

5-2009

Toward Procedural Optionality: Private Ordering of Public Adjudication

Robert J. Rhee

University of Florida Levin College of Law, rhee@law.ufl.edu

Follow this and additional works at: <http://scholarship.law.ufl.edu/facultypub>



Part of the [Legal Profession Commons](#), and the [Litigation Commons](#)

Recommended Citation

Robert J. Rhee, *Toward Procedural Optionality: Private Ordering of Public Adjudication*, 84 NYU. L. Rev. 514 (2009), available at <http://scholarship.law.ufl.edu/facultypub/488>

This Article is brought to you for free and open access by the Faculty Scholarship at UF Law Scholarship Repository. It has been accepted for inclusion in Faculty Publications by an authorized administrator of UF Law Scholarship Repository. For more information, please contact outler@law.ufl.edu.

TOWARD PROCEDURAL OPTIONALITY: PRIVATE ORDERING OF PUBLIC ADJUDICATION

ROBERT J. RHEE*

Private resolution and public adjudication of disputes are commonly seen as discrete, antipodal processes. The essence of private dispute resolution is that the parties can arrange the disputed rights and entitlements per agreement and without judicial intervention. In public adjudication, however, the sovereign mandates the substantive and procedural laws to be applied, many of which cannot be changed by either a party's unilateral decision or both parties' mutual consent. Neither approach allows a party an option to unilaterally alter important aspects of the process, such as the attorney fee rules and standards of proof. This understanding is commonly accepted and rarely challenged, but it is curious nonetheless.

This Article proposes that we move toward procedural optionality, the idea that each party should have options to choose certain procedural laws in public adjudication. To show the potential efficacy of this concept, this Article proposes a scheme in which parties can unilaterally shift fees as long as they contractually bond their good faith by assuming a higher standard of proof. Allowing private choice to alter these rules can better address the problems of frivolous suits and nonprosecution of low value claims—two problematic bookends in the spectrum of litigation. By properly structuring party options, the law can create greater convergence of private incentives and social interest. More efficient dispute resolution results, as measured by increased enforcement of and compliance with the substantive laws, at lower cost. Lastly, this Article concludes by examining more broadly some policy implications of procedural optionality for substantive and procedural laws.

INTRODUCTION	515
I. SOURCES OF SOCIAL COST IN THE LITIGATION SYSTEM.	519
A. <i>The Inflexibility of Inalienable Procedural Rules</i>	519
1. <i>Attorney Fee Rules</i>	520
2. <i>Standards of Proof</i>	521
B. <i>The Social Cost of the Litigation Structure</i>	523
C. <i>Wealth Effects of Fee Rules and Party Incentives</i>	529
1. <i>A Simple Model of Party Incentives</i>	529

* Copyright © 2009 by Robert J. Rhee, Associate Professor of Law, University of Maryland School of Law. J.D., George Washington University Law School; M.B.A., University of Pennsylvania (Wharton); B.A., University of Chicago. I thank my colleague Martha Ertman for our past discussions and Alex Glashauser and Chris Drahozal for their comments. This Article benefitted from comments at the Midwestern Law and Economics Association conference held at Northwestern University School of Law and the Faculty Development Workshop at the University of Maryland School of Law. This Article received generous research support from the University of Maryland School of Law, and I thank Dean Karen Rothenberg. Lastly, I thank my wife Nicki and son Piers for their unconditional support.

2. <i>Party Incentives and Risk</i>	532
II. PROCEDURAL OPTIONALITY OF FEE AND PROOF	
RULES	536
A. <i>Rules of Election</i>	536
B. <i>The Importance of Bonding Good Faith</i>	539
C. <i>Judicial Administration</i>	540
D. <i>Conditions of Election</i>	544
III. SOME IMPLICATIONS OF A BONDED UNILATERAL FEE-SHIFTING REGIME	554
A. <i>Frivolous Lawsuits</i>	554
B. <i>Low Value Cases and Class Actions</i>	557
C. <i>Tort Law as a Case Study</i>	559
IV. THE VALUE OF PRIVATE ORDERING IN PUBLIC ADJUDICATION	562
A. <i>Private Resolution and Public Adjudication</i>	562
B. <i>The Need for Private Ordering in Public Adjudication</i>	565
CONCLUSION	570

INTRODUCTION

The concept of law is one of obligation.¹ This concept has particular force in the realm of procedural and procedure-related laws, such as attorney fee allocation rules and standards of proof. While many ministerial applications of procedural laws are left to the discretion of the parties in public adjudication, the laws are generally not subject to private ordering, either unilaterally or pursuant to agreement. In this sense, certain rules are inalienable and unalterable in public adjudication.² This quality of the adjudicatory framework is axiomatic in our legal culture.

But why should important procedural rules not be subject to a degree of private ordering? At first blush, the suggestion that parties can choose the rules of law seems fanciful. It is commonly thought that the law's ordering of entitlements and obligations is not optional.³ But this proposition overstates the matter. A substantive entitlement in the form of a liability rule, such as tort law, can be seen as an option

¹ See H.L.A. HART, *THE CONCEPT OF LAW* 82 (2d ed. 1994) ("[W]here there is law, there human conduct is made in some sense non-optional or obligatory.").

² A procedural rule is inalienable in the sense that the obligation is not subject to rearrangement by the parties. See Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972) ("An entitlement is inalienable to the extent that its transfer is not permitted between a willing buyer and a willing seller.").

³ See *supra* note 1.

for subsequent reordering through legal pursuit of compensation for the breach of the right; in other words, the defendant has the option to breach the law and pay damages, and the aggrieved plaintiff has the option to prosecute her claim.⁴

More fundamentally for the purposes of this Article, most disputes over legal rights settle without intervention of the public legal process.⁵ While it is said that settlements take place “in the shadow of the law,”⁶ most resolutions are achieved through private ordering rather than public adjudication. Under private ordering, laws can be modified or ignored altogether, should both parties so agree.⁷ But when a dispute enters the realm of public adjudication, our understanding of “law” mostly returns to the model of fixed rules. In judicial proceedings, many procedural laws are inalienable; only courts have the power to rearrange them, and only if the law allows such discretion.

Although this is the generally accepted understanding, it is odd that no one has asked the basic question: Why should certain procedural rules be inflexible, inalienable obligations when so much of dispute resolution is subject to private ordering through such mechanisms as settlements and agreements to arbitrate or mediate? To reject the axiom of fixed procedural laws, we must make only a small, but crucial, conceptual leap. Typically, private ordering and public adjudication are seen as discrete, antipodal processes. There is little intermingling. But if we can envision a hybrid process, one

⁴ Calabresi & Melamed, *supra* note 2, at 1092; see also IAN AYRES, *OPTIONAL LAW: THE STRUCTURE OF LEGAL ENTITLEMENTS* 3–6 (2005) (arguing that substantive entitlements can be analyzed through prism of option theory). Option theory originally arose from the field of financial economics, which developed theories for valuing financial options such as call and put options. See generally RICHARD A. BREALEY ET AL., *PRINCIPLES OF CORPORATE FINANCE* 538–620 (8th ed. 2006) (introducing option theory).

⁵ Most disputes never reach the courthouse because most parties privately agree on the price of the dispute and reach a settlement. See ROBERT C. ELLICKSON, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* 1 (1991) (“A principal finding is that . . . neighbors apply informal norms, rather than formal legal rules, to resolve most of the issues that arise among them.”); H. LAURENCE ROSS, *SETTLED OUT OF COURT: THE SOCIAL PROCESS OF INSURANCE CLAIMS ADJUSTMENTS* 141 (1970) (noting that ninety-five percent of bodily injury claims against insured automobile drivers are settled by negotiation). Even when lawsuits are filed, most cases still settle prior to the final disposition of the matter by the court. See, e.g., Gillian K. Hadfield, *Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases*, 1 J. EMPIRICAL LEGAL STUD. 705, 730 tbl.7 (2004) (finding that settlement rate of contested federal civil cases is approximately sixty-nine percent as of year 2000).

⁶ Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950, 968 (1979).

⁷ See, e.g., ELLICKSON, *supra* note 5, at 65–81 (observing that neighbors often ignore fence laws and apply informal norms that contradict legal prescriptions).

capable of some degree of private ordering of law within public adjudication, then perhaps a party should be allowed to unilaterally change certain important procedural laws or elements of the process. Mutual agreements to alter procedural laws will generally lead parties into alternative dispute resolution, and mutual agreements *not* to alter procedural laws will typically lead parties into court. But what is not currently available in either the public or private dispute resolution forum is an option granted to each party to choose the application of important procedural laws.

The idea of optional laws seems counterintuitive, but it is informed by Ronald Coase's work in *The Problem of Social Cost*.⁸ Coase argued that absent transaction costs, parties can efficiently rearrange rights irrespective of their initial assignment.⁹ Like the substantive entitlements of property and tort law studied in Coase's work, procedural laws set forth an initial assignment of rights and obligations that have significant consequences for the value of lawsuits and their cost to the litigation system. The thesis advanced in this Article is a simple one: The inalienability of certain procedural rules in public adjudication imposes significant cost and risk on parties, while private rearrangement of procedural rights and duties can yield more efficient dispute resolution.

This Article illustrates the efficacy of procedural optionality by analyzing how private ordering of attorney fee rules and standards of proof can mitigate persistent inefficiencies in the litigation system. The problem of inefficient litigation is said to be great and has been the subject of voluminous study and commentary. Two well-known subsets of the larger problem—unpursued, low value, meritorious disputes and frivolous suits—are studied here. Both problems stem from the American rule of attorney fees, according to which each side bears its own fees. In the first subset, high probability, low value disputes are not prosecuted because litigation cost can impose an insurmountable cost barrier to aggrieved persons. This problem has an inverse in the second subset. Here, plaintiffs pursue low probability, frivolous cases to extract an extortionate settlement because defendants sometimes settle to avoid paying a higher cost of defense.

The problems presented by these cases impart significant social cost. They share the same fundamental dynamic: In each case, the probability of success on the merits does not sufficiently influence the ultimate disposition because the attorney fee rule is a fixed legal con-

⁸ R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

⁹ See *id.* at 15 (arguing that assuming no transaction costs, legal rights will be rearranged through market whenever it would increase value).

stant. If the defendant in a frivolous action were able to opt unilaterally for a scheme in which the loser indemnifies the winner's cost, he could avoid settling with the frivolous plaintiff. Likewise, if an aggrieved plaintiff in a low value action could opt unilaterally for fee shifting, she could economically pursue her claim in court. But under the current system of inalienable fee rules, these socially beneficial outcomes are precluded.

The cost structure of the civil litigation system incentivizes frivolous suits and deters low value actions, in direct opposition to the societal interest.¹⁰ When dealing with such procedural defects, the habit of thought has been to propose solutions requiring the imposition of additional inflexible rules. This Article provides a different approach: Rather than intervention in the typical form of rulemaking and judicial application, the sovereign should permit a quantum of private choice in the selection of the procedural scheme in public adjudication. When parties are allowed to reorder certain procedural rules after their initial assignment, dispute resolution can be better achieved, both at the settlement table and in the courtroom.

The specific proposal advanced by this Article is a scheme in which each party can elect to shift its fees to the loser upon prevailing, so long as the party's good faith belief in the merit of the case is bonded by the assumption of a higher standard of proof.¹¹ This scheme would have the greatest effects on the margins of public litigation—claims that are either clearly meritorious or clearly frivolous. This scheme better aligns the results of dispute resolution with important societal interests.

This Article concludes with the broader proposition that we should move toward a greater hybridization of public and private dispute resolution processes. The specific proposal advanced here focuses on frivolous and low value actions, but these problems are not the only inefficiencies in our complex modern litigation system. This Article suggests that future efforts to deal with these inefficiencies consider procedural law not always as inalienable rules but as a set of individual options that, by eliciting the parties' privately held informa-

¹⁰ See Steven Shavell, *The Fundamental Divergence Between the Private and the Social Motive To Use the Legal System*, 26 J. LEGAL STUD. 575, 577 (1997) (arguing that level of litigation is not socially correct because there is divergence between private incentives and social interest).

¹¹ As used in this Article, "to bond" means to secure an obligation by making some form of additional, up-front payment. See BLACK'S LAW DICTIONARY 193 (8th ed. 2004) (defining verb "bond" as "[t]o secure payment by providing a bond"). For example, economists explain the act of purchasing stock as stockholders bonding their contractual obligation to bear the specialized risk of the enterprise by putting up the capital needed to purchase assets. MICHAEL C. JENSEN, *A THEORY OF THE FIRM* 141 (2000).

tion and bonding their good faith belief in that information, can allow for more efficient litigation and settlement.

I

SOURCES OF SOCIAL COST IN THE LITIGATION SYSTEM

A. *The Inflexibility of Inalienable Procedural Rules*

As used in this Article, the term “procedural rules” refers to rules related to the adjudication of substantive rights, including attorney fee rules and standards of proof.¹² No one questions that disputants have significant influence on the administration of public adjudication. They have wide latitude in such matters as scheduling, plan of discovery, and trial presentation, and they routinely petition for favorable application of certain procedural rules. Although these functions are important, they are nevertheless ministerial or supplementary to the court’s power. While smart litigants gain a tactical advantage by navigating these procedural nuances, the most fundamental procedural rules, such as attorney fee rules and standards of proof, are inalienable. True, parties in public adjudication can privately contract for the loser to pay the winner’s legal fees, as is commonly done in arbitration. But these arrangements require a prior relationship and contract, and they do not change the legal rule applicable in court.¹³ Thus, parties cannot contract to alter the attorney fee rules or standards of proof that will be applied by the court in their case.¹⁴ While the concept of inalienable rules in public adjudication

¹² It is acknowledged that courts have defined the standards of proof and attorney fee rules as matters of substantive law. See, e.g., *Marek v. Chesny*, 473 U.S. 1, 35 (1985) (Brennan, J., dissenting) (“The right to attorney’s fees is ‘substantive’ under any reasonable definition of that term.”); *Haberman v. Hartford Ins. Group*, 443 F.3d 1257, 1264 (10th Cir. 2006) (noting that for purposes of diversity jurisdiction, substantive law of forum state governs standard of proof). This Article classifies these rules as procedural because they relate to public adjudication but are essentially unrelated to the underlying entitlement giving rise to the dispute. This term is adopted for simplicity and clarity, and this Article’s argument neither depends on nor intends to provoke a general rethinking of the substance-procedure dichotomy. See generally Jack B. Jacobs, *The Vanishing Substance-Procedure Distinction in Contemporary Corporate Litigation: An Essay*, 41 SUFFOLK U. L. REV. 1 (2007) (providing overview of protean nature of substantive and procedural divide in context of corporate law).

¹³ Additionally, even in arbitration, courts sometimes prohibit the contractual rearrangement of certain entitlements. See, e.g., *DeGaetano v. Smith Barney, Inc.*, 983 F. Supp. 459, 462–63 (S.D.N.Y. 1997) (holding that arbitrator manifestly disregarded law by refusing to award prevailing plaintiff attorney fee under Title VII); *Armendariz v. Found. Health Psychcare Servs., Inc.*, 6 P.3d 669, 682 (Cal. 2000) (holding that arbitration agreement cannot limit statutory remedies such as right to punitive damages and attorney fees).

¹⁴ For example, we do not see a line of employment cases in which the employees have agreed to a higher standard of proof in public adjudication as a condition of the employment contract.

seems to be universally accepted, lack of controversy should not imply correctness. We cannot underestimate the influence of procedural rules on the interests of the parties and society. Since most informed readers are familiar with these rules, only a brief background discussion is needed.

1. Attorney Fee Rules

Different policy considerations support the American and the English rules of attorney fees, and the assignment of a fee rule applies broadly to many classes of cases.¹⁵ Under the English rule, the losing party indemnifies the prevailing party.¹⁶ Under the American rule, each party is responsible for its own legal fees.¹⁷

The most persuasive reason for the English rule is that the prevailing party should be made "financially whole."¹⁸ This argument is compelling in the case of a frivolous suit¹⁹ or in a high probability, low value case (hereinafter simply *low value case*), where the litigation-cost barrier may otherwise bar a highly meritorious claim or defense. But while a loser-pays rule discourages frivolous suits, the effect may be too broad, deterring many risk-averse parties with meritorious claims.²⁰ The English rule discourages low probability, nonfrivolous

¹⁵ See generally Thomas D. Rowe, Jr., *The Legal Theory of Attorney Fee Shifting: A Critical Overview*, 1982 DUKE L.J. 651 (exploring rationales and implications of fee-shifting rules). In addition to these major rules, there is one-way fee shifting, which allows an identified class of disputants, typically plaintiffs, attorney fees upon prevailing. See *id.* at 662–65 (explaining policies of one-way fee rule); see also, e.g., 12 U.S.C. § 4246 (2006) ("When the United States, through private counsel retained under this subchapter, prevails in any civil action, the court, in its discretion, may allow the United States reasonable attorney's fees and other expenses of litigation as part of the costs."); 29 U.S.C. § 216(b) (2006) ("The court in such action shall . . . allow a reasonable attorney's fee to be paid by the defendant, and costs of the action."); *Serrano v. Priest*, 569 P.2d 1303, 1315 (Cal. 1977) (upholding award of attorney fees to plaintiffs under equitable doctrine of private attorney general).

¹⁶ Rowe, *supra* note 15, at 651. But see Geoffrey Woodroffe, *Loser Pays and Conditional Fees—An English Solution?*, 37 WASHBURN L.J. 345, 346 (1998) (noting that small-claims disputes are primary exception in English legal system to general rule of fee shifting).

¹⁷ Rowe, *supra* note 15, at 651.

¹⁸ *Id.* at 657–58.

¹⁹ See Thomas D. Rowe, Jr., *American Law Institute Study on Paths to a "Better Way": Litigation, Alternatives, and Accommodations, Background Paper*, 1989 DUKE L.J. 824, 888 ("Loser-pays fee shifting does have desirable effects—for example, fuller compensation of winners and deterrence of nuisance claims.").

²⁰ See *id.* ("[T]he rule may excessively discourage the pressing of plausible but not clearly winning claims, particularly when the prospective plaintiffs are strongly risk averse."); Thomas D. Rowe, Jr., *Indemnity or Compensation? The Contract with America, Loser-Pays Attorney Fee Shifting, and a One-Way Alternative*, 37 WASHBURN L.J. 317, 330 (1998) (positing that English rule may deter risk-averse claimants from filing meritorious actions).

cases, many of which advance the public interest. The heavy reliance placed on the deterrence argument to justify the fee-shifting rule for *all* classes of cases seems inapposite, given that many cases are neither frivolous nor highly probable. In a close case, a proportional cost allocation may be fairer, and the American rule better approximates this concept.

The policy arguments for and against the American rule are mirror opposites of those advanced in support of the English rule.²¹ The expenditure of attorney fees is not typically considered a legal injury, and fee shifting is only allowed in special classes of cases when justified by separate policy reasons. The American rule encourages more low probability cases, including both frivolous and meritorious actions.²² This promotion of legal actions may be more consistent with American political sentiment, in that our justice system depends on the individual assertion of rights to develop and advance the law. The rule is more egalitarian and better balances the financial disparities that may exist between disputants. Under the English rule, the wealthy person or institution may be privileged to sustain a frivolous or low probability claim, resulting in strategic advantage, while the average person may not.²³

2. *Standards of Proof*

Next, consider the principal standards of proof: beyond a reasonable doubt, clear and convincing evidence (hereinafter the *heightened standard*), and preponderance of the evidence (hereinafter the *default standard*). The standard of proof is a risk-allocation device. In assigning it, the law weighs the risk of error against the interest at stake.²⁴ The beyond a reasonable doubt standard of proof is used in

²¹ The American rule, to the extent that it connotes a general American practice, is somewhat misleading. Exceptions to the rule are commonplace. Rowe, *supra* note 15, at 651 & n.5. There are six primary categories of exceptions: contracts containing fee-shifting provisions, as well as common fund, substantial benefit, contempt, bad faith, and fee-shifting statutes and rules of procedure. John F. Vargo, *The American Rule on Attorney Fee Allocation: The Injured Person's Access to Justice*, 42 AM. U. L. REV. 1567, 1578-90 (1993).

²² See Steven Shavell, *Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs*, 11 J. LEGAL STUD. 55, 59-60 (1982) (noting that frequency of suits is greater under American rule when likelihood of success is below "critical level").

²³ See Rowe, *supra* note 19, at 888 (noting that burden of English rule falls heavily on middle class plaintiffs, who, unlike poor litigants, have money to lose and, unlike rich litigants, cannot afford to lose much).

²⁴ See *Cruzan v. Dir., Mo. Dep't of Health*, 497 U.S. 261, 283 (1990) ("The more stringent the burden of proof a party must bear, the more that party bears the risk of an erroneous decision."). Thus, under the beyond a reasonable doubt standard, the government

criminal law because these cases involve the greatest interests, such as the defendant's life or liberty. The default standard is used in civil litigation because both parties' interests are deemed equally important. Indeed, Coase reformulated this general understanding when he observed that the harm from an activity is typically reciprocal, meaning that often harm to one cannot be prevented or remedied without harm to another.²⁵ This suggests that often the interests of each party in civil cases, unlike criminal cases, cannot be normatively distinguished. The heightened standard is a hybrid of the criminal and civil standards and is usually reserved for an intermediate interest. An example is punitive damages, which are sometimes characterized as "quasi-criminal."²⁶ As a risk-allocation device, the standard of proof, like substantive entitlements, clearly affects the value of a case. The higher the standard of proof, the more difficult it is for a plaintiff to prevail, *ceteris paribus*, and hence the value of her claim decreases.

In the vast majority of cases, the policy debate on the appropriate risk allocation via the standards of proof has long been settled. It is inconceivable that the criminal standard should be anything but the highest, and it is conventionally accepted that the civil standard should allocate the risks equally between plaintiff and defendant. The debate, if there is one, involves determining the interests falling within the heightened standard.²⁷ This bit of untidiness aside, in most cases the application of a particular standard of proof is not a point of contention but a legal constant to be acknowledged.

At some point in a case, the parties must ask the basic question: What are the value, cost, and risk of this case? The primary variables are the set of facts and substantive laws, and perhaps exogenous circumstances such as the draw of the judge, selection of the jurisdiction, attorney quality, and so forth. Some aspects of procedural laws are

bears the substantial burden of erroneous decisions because the criminal defendant's interest at stake is so great.

²⁵ Coase, *supra* note 8, at 2.

²⁶ *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 19 (1991); *James v. Horace Mann Ins. Co.*, 638 S.E.2d 667, 670 (S.C. 2006). Many jurisdictions require punitive damages to be established under the heightened standard. Victor E. Schwartz, Mark A. Behrens & Joseph P. Mastro Simone, *Reining in Punitive Damages "Run Wild": Proposals for Reform by Courts and Legislatures*, 65 BROOK. L. REV. 1003, 1013-14 (1999); see, e.g., OHIO REV. CODE ANN. § 2307.80(A) (West 2004) (establishing standard of awarding punitive damages under "clear and convincing evidence" standard in product liability context).

²⁷ See, e.g., *Conservatorship of Wendland v. Wendland*, 28 P.3d 151, 154 (Cal. 2001) (applying heightened standard to decision to withhold life-sustaining treatment from impaired but conscious person); *In re G.M.*, 596 S.W.2d 846, 846 (Tex. 1980) (applying heightened standard to involuntary parent-child termination proceedings); *State v. Addington*, 588 S.W.2d 569, 569 (Tex. 1979) (applying heightened standard to mental health commitment proceedings).

subject to variable application; for example, the parties have significant input on the process of discovery, and courts have significant discretion in administering the rules of procedure. With that said, procedural rules are for the most part considered legal constants and not variables. Due to legitimate reasons of fairness and predictability, laws such as those governing attorney fee rules and standards of proof make broad distinctions along classes of cases but not finer distinctions based on individual circumstance. In particular, the inalienability of attorney fee rules is regrettable because, ideally, it should be a function of case merit.²⁸

Since the application of these rules is incontestable, much of the assignment of value, cost, and risk is centrally allocated through the sovereign's exclusive rulemaking power. The chief drawback of this approach is its insensitivity to the unique circumstances of each dispute. This not only affects the interests of the parties but also touches upon the important social interests of compliance and enforcement.

B. *The Social Cost of the Litigation Structure*

Assume that a political society deliberates and determines that an activity is undesirable and thus sets a correct "price" for that activity through the provision of rights and remedies.²⁹ Assume further that the concern is not whether the law is normatively correct but only whether the optimal level of compliance and enforcement is achieved. These laws set forth the rules of behavior, society expects compliance, and liability compensates the holder of the right and deters the wrongdoer.

The creation of private rights within the framework of a court system serves two competing interests. First is the goal of perfect compliance, in which the cost of harmful activity is fully internalized by the injurer.³⁰ An alternative utopian ideal is perfect enforcement, which also fully internalizes the cost of one's activity.³¹ Any deviation

²⁸ See *infra* note 57 and accompanying text.

²⁹ See FRANK H. EASTERBROOK & DANIEL R. FISCHEL, *THE ECONOMIC STRUCTURE OF CORPORATE LAW* 37 (1991) (arguing that society can create system of penalties by taxing corporate activity deemed undesirable).

³⁰ Cf. Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347, 348 (1967) ("A primary function of property rights is that of guiding incentives to achieve a greater internalization of externalities.").

³¹ Full internalization does not refer to equivalence in the sense of a market transaction. The suggestion here is that the cost is fully internalized within the limitations of a legal system that compensates injuries, such as lost lives and limbs, with money. See generally Eric A. Posner & Cass R. Sunstein, *Dollars and Death*, 72 U. CHI. L. REV. 537 (2005) (discussing problem of assigning dollar value for life). Such issues have generated enormous controversy. See generally COST-BENEFIT ANALYSIS: LEGAL, ECONOMIC, AND PHILOSOPHICAL PERSPECTIVES (Matthew D. Adler & Eric A. Posner eds., 2001).

from perfect compliance or enforcement creates a social problem. Let us call this problem one of social inefficiency, meaning that some injurers escape paying for the cost of their activities, contrary to society's prescription.³² Because it is clear that perfect compliance is unachievable in the real world, enforcement is the only other means by which to maximize social efficiency.

But enforcement presents significant transaction costs and thus competes with the goal of cost minimization. The prosecution of a private cause of action is costly in the American litigation system, attorney fees being one of the most significant costs.³³ Absent considerations other than compensation (such as the pursuit of justice irrespective of the cost-benefit analysis), it is wasteful to prosecute or defend an action where the cost of doing so exceeds the harm suffered. The interdependencies of social and transaction cost efficiencies suggest that the problem is one of optimization, that is, maximum enforcement and compliance at minimum cost.³⁴

As is evident, transaction costs significantly influence the economics of civil disputes. Conventional law and economics analysis is a helpful starting point for predicting litigation behavior. But first, some terminology must be defined. Each action has a claim for injury, and the estimated value of this injury as would be determined in a judicial forum is hereinafter called the *loss value*. Because the trial outcome in a meritorious action is uncertain, there is a probabilistic

³² Social efficiency must be distinguished from economic efficiency. The concept of social efficiency does not turn on the correctness of a legal rule based on some normative principle but instead refers to the idea that laws should be obeyed and that perfect compliance or enforcement should be the societal goal. Economic efficiency, on the other hand, is defined as the allocation of resources in a way that maximizes value. RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 11 (7th ed. 2007). This suggests that laws maximize economic efficiency. See Richard A. Posner, *A Theory of Negligence*, 1 J. LEGAL STUD. 29, 33 (1972) (arguing that function of negligence liability is to achieve cost-efficient level of accidents and safety). In addition, proponents of economic efficiency propose that the law should be violated if it inefficiently allocates resources. See Frank H. Easterbrook & Daniel R. Fischel, *The Proper Role of a Target's Management in Response to a Tender Offer*, 94 HARV. L. REV. 1161, 1192 (1981) ("If a given act involves a violation of law, penalties may follow. Managers' duty calls for them to minimize the sum of penalty costs and the cost of compliance with the rule.").

³³ Cash costs are primarily attorney fees and litigation costs. In many cases, however, these costs are not the largest cost component in the resolution of a case. See Robert J. Rhee, *The Effect of Risk on Legal Valuation*, 78 U. COLO. L. REV. 193, 198-99 (2007). Valuation discounts given by the parties as a result of the risk profile of the case can be significant. These discounts are not seen as transaction costs because there is no cash expense, unlike attorney fees. Rather, the discount is hidden because it is embedded in the valuation. *Id.* at 199, 229.

³⁴ A classic study of this problem is Gary Becker's analysis of criminal enforcement and sanctions. See generally Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968).

expectation of recovering this loss value. This probability multiplied by the loss value is hereinafter called the *expected value*. Lastly, when each party considers the future payoff or payout from trial, the amount must be net of costs, such as attorney fees. The expected value net of cost is hereinafter called the *net value*.

Conventional analysis predicts that the most important factor in litigation behavior is the ratio of the expected value to the transaction costs of prosecution or defense.³⁵ In a dispute in which the American rule applies, a plaintiff's minimum settlement value is her net value. She will not settle unless the amount offered by the defendant is at least equal to the expected value of going to trial net of cost.³⁶ Similarly, a defendant's maximum settlement value is his net value. A rational defendant would not settle for an amount greater than what he stood to lose at trial, namely the expected value of the claim plus the cost of defense.³⁷ If the defendant's net value at least equals the plaintiff's net value, there is a positive contract zone, and a settlement is possible in the range between their net values. Settlement benefits the parties by creating a surplus from transaction cost savings. But if there is no positive contract zone, the parties cannot agree and trial results.³⁸ Numeric examples illustrate these propositions.

Case 1. The parties agree that the probability of the plaintiff's success is 0.5, the loss value is 100, and each party's transaction cost is 40. The plaintiff's minimum settlement value is the net value of 10. The defendant's maximum settlement value is the net value of 90. The defendant's maximum settlement value is greater than the plaintiff's minimum settlement value. Trial would be wasteful since the parties can settle at the expected value of 50. The spread between their net values is the potential surplus, in this case, a cost savings of 80.

³⁵ See Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 417–18 (1973) (showing how prospect of settlement depends on each party's litigation expenses, settlement costs, and perceived expected value of cost of judgment for plaintiff); Alan E. Friedman, Note, *An Analysis of Settlement*, 22 STAN. L. REV. 67, 80 (1969) (describing parties' bargaining limits as function of expected damage award; probability of award; and litigation, settlement, and opportunity costs). Beyond this basic proposition, there is rich economic and legal literature, too voluminous to cite comprehensively, on bargaining problems.

³⁶ Posner, *supra* note 35, at 417–18.

³⁷ *Id.* at 418; see also *Evans v. Jeff D.*, 475 U.S. 717, 734 (1986) ("Most defendants are unlikely to settle unless the cost of the predicted judgment, discounted by its probability, plus the transaction costs of further litigation, are greater than the cost of the settlement package.").

³⁸ See Robert J. Rhee, *A Price Theory of Legal Bargaining: An Inquiry into the Selection of Settlement and Litigation Under Uncertainty*, 56 EMORY L.J. 619, 631–32 (2006) (discussing conventional law and economics model).

Case 2. Assume that the plaintiff views the probability of prevailing as 0.3. The transaction cost of 40 is greater than the expected value of 30. The plaintiff's minimum settlement value is anything above 0 since the net value is -10. The defendant disagrees with the plaintiff on the merits and believes that the plaintiff's probability of success is 0.1. The defendant's maximum settlement value is 50. A number of possibilities exist for settlement. One possibility is that, with good information on the plaintiff's assessment of her case, the defendant will offer his view of the expected value, which is 10, and thus he will enjoy a surplus of his attorney fee cost savings. Another possibility is that he will offer 5, which is less than his view of the expected value, and thus he will enjoy a rent of 5 derived from the litigation cost structure. The plaintiff may accept these offers because they are better alternatives than trial, which would result in a net value of -10 for the plaintiff, but she must be willing to accept substantial reductions in compensation from her view of the expected value of 30. Alternatively, she can pursue or signal a strategy of brinksmanship and attempt to coerce the defendant into a higher settlement. The eventual outcome is indeterminate and depends on each party's negotiation skills and stomach for risk and gamesmanship.

Case 3. Assume now a frivolous case. The parties believe that the plaintiff's probability of success is close to zero. The plaintiff will settle for any positive value. The defendant's maximum settlement value is his transaction cost of 40. A number of possibilities exist for resolution of this dispute. One is that, as a repeat player concerned about its reputation, the defendant may defend the action to deter future frivolous plaintiffs. Another possibility is that the defendant pays the plaintiff some positive value up to the cost of defense. In this case, the plaintiff earns rent derived from the litigation system.

These simple illustrations show that the social problem of compliance and enforcement can be made more difficult by an expensive litigation system.³⁹ One problem is that when expected value is less than the transaction costs of pursuing an action to trial, the prosecution of an action is uneconomical—why pursue a claim for one dollar when it costs two dollars to litigate it? The cost structure of litigation can preclude the prosecution of low value cases.⁴⁰ Absent a positive

³⁹ Cf. Samuel R. Gross & Kent D. Syverud, *Don't Try: Civil Jury Verdicts in a System Geared to Settlement*, 44 UCLA L. REV. 1, 62 (1996) ("We have designed a spectacular system for adjudicating disputes, but it is too expensive to use.").

⁴⁰ To be clear, "low value" is not an absolute concept of some threshold dollar value but rather refers to the ratio of the expected value to transaction costs. For example, a case for injury of \$100,000—a substantial sum—qualifies as a low value case if the expected litigation cost is \$90,000.

settlement, a rational plaintiff may resolve a low value case by waiving her right to recovery, though the law may provide a *de jure* opportunity for remedy.⁴¹ This in effect gives the defendant—the wrongdoer—substantial power to set his level of liability.⁴² For these types of cases, merit is often dissociated from the decision to litigate a claim.

This does not mean that holders of low value cases are without access to justice. Procedural mechanisms such as class action and consolidation were created to redress this persistent problem.⁴³ But access to justice through these mechanisms depends less on the merit of any individual claim and more on the fortuity of finding an action or similarly situated persons who can share the cost and risk. Many potentially meritorious class or consolidated actions go unprosecuted because of the high cost of searching out and organizing these actions. Absent good fortune, the cost structure of the civil litigation system excludes many cases based solely on a financial ratio.

Another problem is that the existing cost structure creates the condition for frivolous litigation.⁴⁴ In a world of zero transaction costs

⁴¹ See Rowe, *supra* note 19, at 889 (discussing expense of litigation as barrier to full litigation of small claims). Scholars have noted that many meritorious claims for injuries are not prosecuted. DEBORAH R. HENSLER ET AL., RAND, COMPENSATION FOR ACCIDENTAL INJURIES IN THE UNITED STATES 142 (1991) (“[C]laiming is a statistically *unusual* behavior: many more injured people decline to claim—or never even consider claiming—than attempt to activate the legal process.”); see also TOM BAKER, THE MEDICAL MALPRACTICE MYTH 68–69 (2005) (noting that many meritorious medical malpractice claims go unprosecuted); Marc A. Franklin et al., *Accidents, Money, and the Law: A Study of the Economics of Personal Injury Litigation*, 61 COLUM. L. REV. 1, 10 (1961) (observing that most claims for injury in New York are never adjudicated in court); Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System—and Why Not?*, 140 U. PA. L. REV. 1147, 1183 (1992) (“A great many potential plaintiffs are never heard from by the injurers or their insurers.”). This empirical fact obviously undermines the notion that a litigation compensation scheme such as tort law is economically efficient. See Robert J. Rhee, *Tort Arbitrage*, 60 FLA. L. REV. 125, 129–30, 167–69 (2008) (arguing that theory of negligence cannot be economically efficient because of nonprosecution of injuries and undervaluation of plaintiff claims).

⁴² This is the precise situation in which a corporation commits low value infractions against consumers, which makes individual lawsuits infeasible without fee shifting. Left unprosecuted, the aggregate benefit from such repeated violations can result in great financial gain to the wrongdoer. The result is that rent can be earned from the litigation cost structure, or the wrongdoer can select the level of its liability vis-à-vis society through the political process.

⁴³ See FED. R. CIV. P. 19 (required joinder of parties); FED. R. CIV. P. 20 (permissive joinder of parties); FED. R. CIV. P. 23 (class actions); see also *infra* note 128 (providing rationale for class actions).

⁴⁴ As the term is used here, a frivolous suit has three essential elements: (1) The plaintiff knows or should know that the action has no merit; (2) the probability of success is in fact very small; and (3) the plaintiff’s motive is to procure a positive settlement value. This is the core of a Rule 11 inquiry. See FED. R. CIV. P. 11(b) (providing that by presenting papers to court, attorney certifies that filing is not presented for improper purposes and is supported by nonfrivolous legal arguments and factual contentions). A frivolous lawsuit

(aside from one's own effort), the only frivolous suits remaining would result from either gross mistake or a large difference in the opportunity cost of effort between the parties such that the defendant would pay to avoid this cost.⁴⁵ But our current world of high transaction costs enables the logic of frivolous litigation—a logic of extortion and brinksmanship. A frivolous suit is valueless absent a credible threat of harmful cost imposition from prosecution,⁴⁶ and a determined plaintiff can make such a threat, thereby coercing a defendant to pay for the avoidance of greater financial harm. If a defendant is not subject to repeat play consideration, he may rationally settle if settlement is cheaper than litigation. Thus, the cost structure of the legal system not only deters the pursuit of meritorious low value cases, it also provides no public incentive to defend frivolous actions outside of Rule 11 sanction, which is often cold comfort given that it requires adjudication on merit and motive.⁴⁷

In summary, cost drives many of the decisions and strategies concerning prosecution and defense in a legal system. The problems of unprosecuted low value disputes and frivolous suits can be seen as bookends of the litigation spectrum. These cases are linked by a similar (but inverse) probabilistic profile, and both are governed by the same dynamics of transaction costs. Each problem creates a different form of social cost: underprosecution in low value cases and overprosecution in frivolous suits. This result runs precisely opposite to the societal interest since, ideally (disregarding the utopian possibility of perfect compliance), all frivolous suits should be deterred at

can also be brought to harass or annoy a party. Here, the problem of transaction costs may be more peripheral; anyone willing to harass or annoy another at the cost of one's own attorney fees under the American rule may still do so under the English system if the ill will is strong enough.

⁴⁵ In a world of zero transaction costs, a great many frivolous suits might still be filed because there is always the chance that the court might make a mistake and because the opportunity to litigate would be cost free. But in such a system, the opportunity cost of one's effort provides the minimum cost barrier, which could reduce the number of frivolous suits to near zero if the plaintiff's opportunity costs are greater than the expected value of a mistake by the court.

⁴⁶ See Lucian Arye Bebchuk, *A New Theory Concerning the Credibility and Success of Threats To Sue*, 25 J. LEGAL STUD. 1, 2 (1996) (noting that defendant will not agree to settle negative-net-value suits unless plaintiff's threat to litigate is credible). Consider also that insurers often pay a "danger value," a premium in excess of the risk-neutral value, on low probability claims involving death or serious injury. This is not because the insurer is likely to be found liable but because the insurer seeks to avoid the possibility, however slight, of high litigation payouts. Ross, *supra* note 5, at 199–204.

⁴⁷ See *infra* notes 175–78 and accompanying text (discussing problems with overreliance on Rule 11 and similar court-imposed sanctions as means for deterring frivolous suits).

inception or defended on the merits, and all meritorious actions should be prosecuted.⁴⁸

C. *Wealth Effects of Fee Rules and Party Incentives*

I. *A Simple Model of Party Incentives*

The effects of procedural rules on the value, cost, and risk of a case deserve closer scrutiny. It is useful to set forth a simplified model of two-party dispute resolution, as it provides a frame of reference for subsequent discussion of attorney fee rules. For the remainder of this Article, the following notations and value assignments are used:

V = each party's view of net value

P = each party's view of probability of liability

L = loss value ($L = 100$)

T = each party's transaction costs ($T = 20$)

C = aggregate transaction costs ($C = 2T = 40$)

For simplicity, the loss value and transaction costs are fixed such that $L = 100$, $T = 20$, $C = 40$ (see above). Holding these values constant results in binary trial outcomes $\{100, 0\}$. This simplification reduces the valational variables to two: (1) each party's probability assessment (and any rule that affects this calculation, such as the standard of proof), and (2) the law's assignment of the fee rule.

Under the American rule, the two probabilistic outcomes are a win or a loss, both of which are net of one's transaction costs. Under the English rule, the outcomes are a win with fee indemnity or a loss net of aggregate cost assignment. Under a one-way rule in which the party stands to benefit from fee shifting (obviously one can also be on the adverse side of a one-way fee shift), the outcomes are a win with fee indemnity or a loss net of one's transaction costs. The fee rules can be formalized into the following payoffs and payouts under the American and English rules and a beneficial one-way rule of fee shifting, as laid out in Table 1 below. When discussing the defendant's payoffs and payouts, we also assume that defendants can enjoy the benefit of one-way fee shifting upon prevailing, even though most actually implemented one-way fee-shifting schemes are designed to benefit plaintiffs rather than defendants.⁴⁹

⁴⁸ See *supra* notes 29–34 and accompanying text (discussing problems of social cost, compliance, and enforcement).

⁴⁹ See *supra* note 15.

TABLE 1

	American	English	One-Way
Plaintiff	$V_p = PL - T$	$V_p = P(L + C) - C$	$V_p = P(L + T) - T$
Defendant	$V_d = PL + T$	$V_d = P(L + C)$	$V_d = P(L + T)$

Without exception, a party favors a one-way rule to his benefit and disfavors its reverse application. The application of the American and English rules is symmetric, and a party's preference between these two fee rules depends on her probability of prevailing. If the probability of prevailing is greater, a plaintiff would prefer the English rule over the American rule. The plaintiff's net value must be greater under the English rule, since fee recovery is greater than fee liability if the probability of success is more likely than not.⁵⁰ A numeric example illustrates the point. Assume that $P = 0.6$. The parties' net values are:

$$\begin{array}{ll} \text{Plaintiff} & \left\{ \begin{array}{ll} \text{American} & 40 = PL - T = 0.6 \times 100 - 20 \\ \text{English} & 44 = P(L + C) - C = 0.6(100 + 40) - 40 \end{array} \right. \\ \\ \text{Defendant} & \left\{ \begin{array}{ll} \text{American} & 80 = PL + T = 0.6 \times 100 + 20 \\ \text{English} & 84 = P(L + C) = 0.6(100 + 40) \end{array} \right. \end{array}$$

The English rule increases the plaintiff's net value from 40 to 44. Concomitantly, the defendant's net value decreases from 80 to 84 (a decrease in net value because liability increases).

The opposite must also hold. If the probability of success is less likely than not, the plaintiff prefers the American rule.⁵¹ For example, assume that $P = 0.4$, and apply the above calculations. The plaintiff's minimum settlement values are {20 American, 16 English}, and the defendant's maximum settlement values are {60 American, 56 English}.⁵²

In addition to changing net values, the English rule also produces greater variance of expected outcomes. Under the American rule, the plaintiff's binary outcomes of net value are {80, -20}, either a win or a loss net of one's fees. The total spread of outcomes is 100, which is the loss value at stake. Under the English rule, the plaintiff's binary outcomes of net value are {100, -40}. The total spread is 140, which is

⁵⁰ We can show this algebraically: If $P > 1 - P$, meaning $P > 0.5$, then it must be true that $P(L + C) - C > PL - T$. This inequality reduces to $PC > T$. Since $P > 0.5$, $PC > T$ if $C = 2T$.

⁵¹ If $P < 1 - P$, meaning $P < 0.5$, then it must be true that $P(L + C) - C < PL - T$. This inequality reduces to $PC < T$ if $C = 2T$.

⁵² The plaintiff's calculations are as follows: $PL - T = 0.4 \times 100 - 20 = 20$ (American rule); $P(L + C) - C = 0.4(100 + 40) - 40 = 16$ (English rule). The defendant's calculations are as follows: $PL + T = 0.4 \times 100 + 20 = 60$ (American rule); $P(L + C) = 0.4(100 + 40) = 56$ (English rule).

the loss value *plus* the aggregate cost.⁵³ The preference for a particular fee rule is a function of the probability assessment, yielding different cost and net value. But the English rule always increases the variance of outcomes (the potential deviation of actual result from the probabilistic expectation)⁵⁴ and thus the riskiness of a case.⁵⁵

To summarize, the assignment of the fee rule has effects on each party's wealth that are similar to the effects of substantive entitlements. A plaintiff with a low probability or frivolous case would prefer the American rule.⁵⁶ Conversely, the party with a high probability case would prefer the English rule. These clear preferences suggest certain policy prescriptions. Few would dispute that the frivolous plaintiff should lose and be made to indemnify the cost of the defendant. There is also a compelling case for fee shifting in favor of a plaintiff with a very high probability case—one against a marginally frivolous defendant, perhaps. In the ideal world, then, an important consideration in the assignment of the fee rule is *case merit*.⁵⁷ Thus far, we have assumed that this can only be done in public adjudication through the decision of the court.⁵⁸ Outside of judicially imposed sanctions such as Rule 11, which are aimed at abusive behavior, the debate over the assignment of fee rules has focused pri-

⁵³ Richard L. Schmalbeck & Gary Myers, *A Policy Analysis of Fee-Shifting Rules Under the Internal Revenue Code*, 1986 DUKE L.J. 970, 977 ("[T]he clear effect of using the English rule . . . is to increase each party's stakes by the sum of *both* parties' litigation costs, not merely by the other party's costs.").

⁵⁴ See Rhee, *supra* note 33, at 199 ("Thus, variance is defined as the measure of one's belief about the possible deviations of a judgment from expectation, and it gauges the subjective perception of uncertainty.").

⁵⁵ Only when the probabilities are in equipoise does the expected value remain the same under the American and the English rules. If $P = 1 - P$, meaning $P = 0.5$, the plaintiff's returns are the same: $P(L + C) - C = PL - T$. The same is true for the defendant's returns: $P(L + C) = PL + T$.

⁵⁶ See James W. Hughes & Edward A. Snyder, *Litigation and Settlement Under the English and American Rules: Theory and Evidence*, 38 J.L. & ECON. 225, 249 (1995) ("The English rule likely reduces the frequency of low-merit claims . . ."); David Rosenberg & Steven Shavell, *A Model in Which Suits Are Brought for Their Nuisance Value*, 5 INT'L REV. L. & ECON. 3, 5 (1985) ("[U]nder the British system the willingness of the plaintiff to litigate and to file a claim will be less than under the . . . American system if the likelihood of prevailing is low.").

⁵⁷ See Lucian Arye Bebchuk & Howard F. Chang, *An Analysis of Fee Shifting Based on the Margin of Victory: On Frivolous Suits, Meritorious Suits, and the Role of Rule 11*, 25 J. LEGAL STUD. 371, 372, 374 (1996) (arguing that fee shifting should be based on adjudicated case merit).

⁵⁸ See, e.g., *id.* (requiring adjudication and judicial discretion in proposal to shift fees based on case merit). But see *infra* notes 175–78 and accompanying text (critiquing this aspect of Bebchuk and Chang's proposal).

marily on distinguishing classes of cases based on policy considerations other than case merit.⁵⁹

A last point is warranted. The economic model employed here is highly stylized, and no serious person believes that dispute resolution is so rationally programmatic. The empirical world is messy and ambiguous. The economic model assumes information symmetry and risk neutrality, meaning that the parties have the same information and care only about expected return. It thus ignores both the potential range of deviation from expectation and the uncertainty arising from imperfect information or process, which together define litigation risk.⁶⁰ Risk is dismissed because conventional analysis assumes the rational expectation model, that is, that "the subjective expectations correspond to the objective frequencies of the random event."⁶¹ Of course, this model is unrealistic since "uncertainty is *the* essential characteristic of a lawsuit," and perhaps the principal motivator of litigation behavior.⁶²

2. *Party Incentives and Risk*

The attorney fee rules affect not only the value and cost of an action but also its risk. The English rule increases the variance of outcomes. The assumption of risk neutrality as a predictive matter is facially absurd if one considers how parties actually behave, not only in the context of litigation, but also in markets in general. Disputants and market participants care greatly about risk. The capital and insurance markets, for example, exist precisely because people have dif-

⁵⁹ It is of course true that even if a fee rule is made to apply to a certain class of cases regardless of individual case merit *within* that class, prevailing on the merits is a condition of a fee award.

⁶⁰ In the context of litigation analysis, risk is defined as the degree of variance from one's expectation of the outcome, and variance is typically measured by one's confidence in one's assessment of the case. Rhee, *supra* note 33, at 199, 217-19; Rhee, *supra* note 38, at 679.

⁶¹ Robert Cooter, *The Cost of Coase*, 11 J. LEGAL STUD. 1, 22 (1982); see also George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 9 (1984) (assuming existence of "true value" of litigation). Probability in legal assessment is not a numeric, objective measure. It is a subjective, logical concept, allowing for differences in "ex ante degrees of rational belief in the plausibility of these decision paths, irrespective of the ex post result." Rhee, *supra* note 38, at 653. I do not yet know a lawyer, judge, or scholar who can consistently predict the outcomes of trials, and if a person were to possess such a unique skill, she would certainly not be a lawyer, judge, or scholar for long, as she would have retired long ago to her private island.

⁶² Rhee, *supra* note 33, at 215 (emphasis added); see also Rhee, *supra* note 38, at 665 ("The nature of litigation is stochastic . . .").

ferent views not only of the expected value of future income but also of the risk associated with that income.⁶³

We add complexity (and a dose of reality) by considering the effect of risk on behavior. Market participants prefer to maximize return and minimize risk. It is foolish to take unnecessary risk.⁶⁴ Disputants in the litigation system also attempt to maximize return and minimize risk.⁶⁵ Risk management is common in capital markets, insurance transactions, and gaming, and there is no reason to believe that the benefits of risk management do not extend to litigation as well. In a risk-based model of dispute resolution, disputants are seen as constructing implied insurance/derivative transactions where risk and return are traded.⁶⁶ Thus, for example, when two disputants share the same view of expected value but hold different opinions of the range of variance, the party perceiving greater risk would value the litigation asset less (and conversely the liability more) than the other party when deciding whether to settle or litigate. He therefore trades risk for a valuational concession in expected value, and this implied hedging process creates divergent views on the value of the asset/liability of the disputed right.⁶⁷

⁶³ See FRANK H. KNIGHT, *RISK, UNCERTAINTY AND PROFIT* 268–71 (Dover 2006) (1921) (noting that enterprise and markets exist because of uncertainty); see also BREALEY ET AL., *supra* note 4, at 188–91 (discussing relationship between risk and return); SCOTT E. HARRINGTON & GREGORY R. NIEHAUS, *RISK MANAGEMENT AND INSURANCE* 171–73 (2d ed. 2004) (discussing why corporations purchase insurance).

⁶⁴ For instance, diversification, a bedrock principle of financial economics, is preferred because it obeys the rule that “the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing.” Harry Markowitz, *Portfolio Selection*, 7 J. FIN. 77, 77 (1952).

⁶⁵ See Rhee, *supra* note 33, at 237 (discussing relationship of risk and expected value); see also J.B. Heaton, *Settlement Pressure*, 25 INT’L REV. L. & ECON. 264, 272 (2005) (noting that even if corporations could be presumed risk neutral because they are non-human entities, they can still behave in ways that reflect risk aversion).

⁶⁶ See Rhee, *supra* note 33, at 229–39 (providing detailed analysis of this trading mechanism between parties).

⁶⁷ A simple example illustrates this concept. Assume that trial outcome is binary. The plaintiff believes that the outcomes are equally probable {100, 0}, that is, a win of 100 or a loss yielding 0. The expected value is 50. The defendant has a different view of the binary outcomes {25, 75}, that is, a guaranteed loss but of either 25 or 75. The defendant shares the plaintiff’s assessment of the expected value of 50. Without considering risk, we expect that the parties would settle at 50.

However, if the valuational framework takes risk into account, as is the case in the financial and insurance markets, then the value is adjusted during the settlement process. Clearly, the plaintiff believes the outcomes to be riskier, since there is greater variance in her assessment. To mitigate risk, the parties will implicitly issue put options (or insurance policies) to one another, providing, in essence, side payments in the event that the other party experiences an adverse event. In the defendant’s case, he expects to pay out the expected value of 50, but this value is subject to variation of up to 25. The insurance policy thus has a face value of 25. The plaintiff issues the policy and in consideration receives a

A risk-based analysis is relevant to the controversy concerning the impact of the fee rule on litigation and settlement.⁶⁸ Since the English rule enhances net value if the probability of prevailing is greater, some have argued that it promotes more trials.⁶⁹ Others have argued the opposite, reasoning that the possibility of an adverse fee shift could result in a greater tendency to settle under the English rule.⁷⁰ Still others have argued that there is no clear discernable effect.⁷¹

Without consideration of risk, the two rules should have the same effect on trial and settlement. The principle of conservation of value applies here.⁷² Each party's net value may change with a change in

reciprocal policy of 25 from the defendant. But this policy only reduces the variance of outcomes from {100, 0} to {75, 25}. There is still a residual risk.

To mitigate the risk further, the plaintiff must seek additional insurance from the defendant. The defendant will issue an additional policy (that is, agree to settle, thus terminating the exposure to risk) but will do so only for additional consideration. The plaintiff must now give up some expected value for the issuance of additional insurance or a put option by the defendant. Thus, the settlement value may and probably will, in many cases, be less than the expected value. *See id.* at 233–37 (providing formal model of risk-adjusted litigation valuation).

⁶⁸ *See* POSNER, *supra* note 32, at 618 (“The most hotly debated question about [fee] indemnity is its effect on the litigation rate, with its advocates touting indemnity as the answer to the caseload crisis.”).

⁶⁹ *See* A. Mitchell Polinsky & Daniel L. Rubinfeld, *Does the English Rule Discourage Low-Probability-of-Prevailing Plaintiffs?*, 27 J. LEGAL STUD. 141, 143 (1998) (“[T]he English rule causes a greater number of cases to go to trial, and all of these additional cases involve plaintiffs whose probability of prevailing is less than that of plaintiffs who go to trial under the American rule.”); Jeffrey J. Rachlinski, *Gains, Losses, and the Psychology of Litigation*, 70 S. CAL. L. REV. 113, 161 (1996) (“[B]y raising the stakes at trial, the loser-pays system makes litigation itself more valuable and can discourage settlement.”); Shavell, *supra* note 22, at 65 (“[T]he likelihood of trial under the British system will be greater than under the American system.” (emphasis omitted)); *cf.* POSNER, *supra* note 32, at 618 (arguing that English rule promotes more trials when party believes it will prevail); Keith N. Hylton, *Fee Shifting and Incentives To Comply with the Law*, 46 VAND. L. REV. 1069, 1079 (1993) (“[T]he British rule raises the stakes, which makes litigation more attractive to the parties when the plaintiff places a higher estimate on the likelihood of his winning than does the defendant.”).

⁷⁰ *See* Thomas D. Rowe, Jr., *Predicting the Effects of Attorney Fee Shifting*, 47 LAW & CONTEMP. PROBS. 139, 159 (1984) (“[A]dding the possibility of a fee shift against individual litigants relying on their own resources might well result in a greater tendency to settle claims once pursued than exists under the American rule.”).

⁷¹ *See* Schmalbeck & Myers, *supra* note 53, at 976–77 (“[T]he incentive and disincentive effects of the English rule would seem, *a priori*, to be in rough equipoise.”).

⁷² This term is borrowed from the financial economic principle that the value of an asset is preserved regardless of the nature of the claims against it (that is, the capital structure of the firm). BREALEY ET AL., *supra* note 4, at 448. This principle was announced in seminal articles by Franco Modigliani and Merton Miller. *See generally* Franco Modigliani & Merton H. Miller, *The Costs of Capital, Corporation Finance and the Theory of Investment*, 48 AM. ECON. REV. 261 (1958); Franco Modigliani & Merton H. Miller, *Corporate Income Taxes and the Cost of Capital: A Correction*, 53 AM. ECON. REV. 433 (1963). This principle states that, absent bankruptcy costs and interest expense tax shields, the capital

the fee rule (in other words, the nature of each party's claim to the economic pie—loss value plus aggregate cost—may change), but the size of this pie remains the same in the sense that the fee rules do not change the loss value or the contractual rate of the fees. The only question is whether the rules affect the timing of settlement and thus the amount of fees expended by promoting either settlement or trial. However, since the rules affect only the *allocation* of fees, the increased incentive to litigate for one party should be offset exactly by an increased incentive to settle for the other party.

A numeric example illustrates this principle. In the above example where $P = 0.6$, we calculated the parties' values under the American rule as {40 plaintiff, 80 defendant} and their respective values under the English rule as {44 plaintiff, 84 defendant}. The contract zone under both rules is 40, which constitutes the surplus of aggregate costs saved by settling rather than proceeding to trial. The plaintiff's increased net value under the English rule (from 40 to 44) is exactly offset by the defendant's willingness to pay an increased settlement (from 80 to 84). This implies that, barring other factors, attorney fee rules should be irrelevant.

This conclusion, however, does not consider the effect of risk. The English rule produces more variable returns. If one assumes that most people are risk averse and strive to mitigate risk exposure,⁷³ the English rule may work at the margin to systematically push cases to settlement. Since greater variance of outcome produces greater opportunities to trade risk and return, we expect that disputants will actively engage in risk mitigation measures, resulting in more opportunities for settlement. Importantly, the risk profiles of parties in any given case typically will differ (that is, some parties are better financed, are wealthier, are less risk averse, or can diversify risk better), and the party who has an advantage in this regard under the circumstances will gain a valuation advantage and will extract a price concession from the other party in settlement.⁷⁴

structure cannot change the value of a firm's output. Stated differently, reconfiguration of the capital structure cannot augment or diminish the value of the underlying economic activity.

This principle is relevant in the litigation context. The rules governing attorney fees do not change the loss value (the underlying liability), the contractual rate of fees, or the amount of the attorney fees expended. The rules simply determine how cost is allocated between the parties.

⁷³ See *Price v. Marshall Erdman & Assocs., Inc.*, 966 F.2d 320, 327 (7th Cir. 1992) (Posner, J.) ("[M]ost people are assumed to be risk-averse in their serious financial affairs . . .").

⁷⁴ Rhee, *supra* note 41, at 181.

In summary, the purpose here is not to suggest a definitive answer to the controversy concerning which rule promotes settlement. We can make educated guesses supported by theoretical arguments, but ultimately the controversy probably will be laid to rest by empirical evidence.⁷⁵ The point made here is the modest proposition that assessing the value of a lawsuit requires a multivariate analysis. The wealth effects of procedural rules influence the assignment of not only cost but also risk, both of which affect value and strategy. Given the fee rule's influence, how should one apply it so as to advance the legitimate interests of the parties and society?

II

PROCEDURAL OPTIONALITY OF FEE AND PROOF RULES

A. *Rules of Election*

Given how attorney fee rules and standards of proof can affect the net value of a case and the incentives of the parties, it is puzzling that courts and lawmakers have not attempted to link them. Outside of sanctions, fee rules are assigned based on the taxonomical organization of the policies underlying the cause of action; the standard of proof has been based on the singular consideration of weight of interest.⁷⁶ The static nature of the assignment function means that currently the rules are not efficacious mechanisms for shaping incentives and closing the divergence of private and social interests. This is regrettable since procedural rules are fundamentally risk- and cost-allocation devices that the parties can use to negotiate more efficient dispute resolutions.

The task should be to implement an incentive and deterrence structure that optimizes enforcement and cost. To address this problem and, more broadly, to demonstrate the efficacy of the concept of procedural optionality, this Article proposes a scheme in which each party has the power to unilaterally shift fees to the other side as long as she bonds her good faith belief by assuming a higher standard of proof. At the outset, it is important to state the baseline assumptions. Under the proposal advanced here, the default state is assumed to be the American rule of attorney fees and the preponderance of the evidence burden of proof. The proposal here does not consider situa-

⁷⁵ With regard to empirical evidence, differences in the legal systems and cultural and political mores contribute to the litigation habits of those subject to the American and English rules. Rowe observed that contrary attitudes explain the difference in the rules: "The English seem to begin from an attitude that losers should naturally pay, while Americans tend to want affirmative justifications to support fee shifts." Rowe, *supra* note 15, at 654 n.13.

⁷⁶ See *supra* Part I.A (providing background on fee rules and standards of proof).

tions outside of these parameters, such as claims brought under existing statutes providing for fee shifting. Additionally, this Article should not be construed to promote a specific legislative agenda or cause but instead to provide a conceptual framework for a more flexible procedural regime, which can then be considered in the context of substantive rules and policies.

With that said, the rules of election proposed by this Article are simple. In a two-party suit, there are two standards of proof and two attorney fee rules available, leaving four possible combination scenarios.

Scenario 1: Neither Party Elects. Absent some other laws, the status quo remains, that is, the default standard of proof and the American rule apply.

Scenario 2: Both Parties Elect. Both the default and heightened standards of proof apply. Where either party proves its case under the default standard, the American rule applies. The election of the heightened standard becomes a nullity because both parties failed to meet it. The consequence of bilateral fee shifting occurs only when one party prevails under the heightened standard, in which case fees are shifted to the loser. As discussed later in this Article, this scenario should almost never occur once the parties have exchanged some information through discovery.⁷⁷

Scenario 3: Plaintiff Elects. The plaintiff unilaterally imposes one-way fee shifting on the defendant, but she can only win by proving her case under the heightened standard, upon which fees are shifted to the defendant. This means that if she proves her case under the default standard but fails to meet the heightened standard, she will lose, though the defendant will not be entitled to fee shifting.

Scenario 4: Defendant Elects. The defendant unilaterally imposes one-way fee shifting on the plaintiff, but he can only win by proving his case under the heightened standard, upon which his fees are shifted to the plaintiff. This means that if he fails to meet the heightened standard, the plaintiff will win even if she fails to prove her case under the default standard, though she will not be entitled to fee shifting.

These simple rules of election and bonding allow the parties to reorder the value, cost, and risk of a case based upon the incentives of each party to maximize enforcement of a right at the lowest cost. The following table summarizes these election rules and their consequences.

⁷⁷ See *infra* Part II.D.

TABLE 2

Did the Party Elect the Heightened Standard?		→	Can the Party Win by Satisfying the Default Standard? (No Fee Recovery)		→	Can the Party Collect Fees upon Winning Under the Heightened Standard?	
Plaintiff	Defendant		Plaintiff	Defendant		Plaintiff	Defendant
No	No	→	Yes	Yes	→	No	No
Yes	Yes	→	Yes	Yes	→	Yes	Yes
Yes	No	→	No	Yes	→	Yes	No
No	Yes	→	Yes	No	→	No	Yes

These election rules provide a degree of symmetry in risk and return, but the symmetry should not be overstated. While the rules are symmetrical, the inherent roles of a plaintiff and a defendant are different. Unless there is a legal presumption of some sort, such as the doctrine of *res ipsa loquitur*, the plaintiff must present affirmative evidence of each element of the cause of action.⁷⁸ Because the plaintiff bears this obligation, the defendant can prevail under different circumstances: The plaintiff can fail to present a *prima facie* case, which would typically result in a summary judgment;⁷⁹ the defendant can negate an element of the plaintiff's *prima facie* case;⁸⁰ or the defendant can prove an affirmative defense. The differences in the criteria for prevailing ensure a certain degree of asymmetric risk and burden.⁸¹ Accepting these differences, the above rules strive for symmetry of application, which is essential for fairness and acceptability.

⁷⁸ See, e.g., RESTATEMENT (SECOND) OF TORTS § 328A (1965) (setting forth burden of proof in negligence actions).

⁷⁹ See *Celotex Corp. v. Catrett*, 477 U.S. 317, 322 (1986) (finding that summary judgment is warranted when party "fails to make a showing sufficient to establish the existence of an element essential to that party's case, and on which that party will bear the burden of proof at trial").

⁸⁰ See *Adickes v. S.H. Kress & Co.*, 398 U.S. 144, 157 (1970) (holding that defendant moving for summary judgment must negate nonmoving plaintiff's averment of facts).

⁸¹ Clearly, the substantive law can change these risk and burden allocations. For example, burden shifting in tort law is common when policies call for the reallocation of the risk of error. See, e.g., *Sindell v. Abbott Labs.*, 607 P.2d 924, 937 (Cal. 1980) (shifting burden of proof on factual causation to defendant under theory of market share liability); *Haft v. Lone Palm Hotel*, 478 P.2d 465, 475 (Cal. 1970) (holding that "the burden of proof on the issue of causation should be shifted to defendants to absolve themselves if they can" when defendants failed to observe statute on swimming pool safety); RESTATEMENT (SECOND) OF TORTS § 328D cmt. m (1965) ("In the ordinary case the great majority of the courts now treat *res ipsa loquitur* as creating nothing more than a permissible inference, which the jury may draw or refuse to draw, unless the facts are so compelling that no reasonable man could reject it."); see also Robert J. Rhee, *Probability, Policy and the Problem of Reference Class*, 11 INT'L J. EVIDENCE & PROOF 286, 290-91 (2007) (arguing that "systemic error" from interaction of procedural and substantive laws can require "normative correction" in substantive law).

B. *The Importance of Bonding Good Faith*

If public adjudication is to rely on the parties' input to determine the degree of merit, and thus which fee rule to apply, we need a way to penalize bad faith or failed performance. A scheme of optional laws would quickly unravel into an anarchist's impasse if there were no mechanism by which to bond the parties' good faith.⁸² The bond must be priced correctly, meaning that the penalty should be proportional to the benefit of the option. If it costs too much, the option will be uneconomical and the status quo will rule. If it costs too little, the option will be exercised automatically, and a wholesale rearrangement of the initial assignments of the procedural rules will result. For a scheme of optional laws to work, there must be proportionality between incentive and disincentive such that the effects of self-interest and strategic behavior serve a normative social goal, or at least minimally conflict with it.

We can derive a set of bargaining rules by considering the scheme of procedural election the parties would produce if they were ignorant of their role in a subsequent litigation. We must ask: Would parties, uncertain of their status as disputants, prefer to submit to an inflexible rule or the discretion of the court, or would they prefer to choose the application of rules themselves in each individual case? For most people, control of the variables of outcome affecting one's interest is a good thing, particularly when there is little or no information on the tendencies of the alternative adjudicator and process. Accepting this premise, we must ask further: Given the problem of self-interest and bonding, what would those rules of election be in a hypothetical bargain?

The bonding aspect of procedural optionality facilitates better conveyance of information. The act of election conveys the information that the electing party has significant information, is confident of the completeness of the information such that variance is not a large concern, and in good faith believes in his probabilistic assessment of success. The value of this information communicated is high because it is reliable. It is reliable because a party voluntarily imposes a substantial cost on himself to communicate it, namely the assumption of a higher standard of proof.

⁸² This is the reason why substantive laws cannot be subject to election, for it is difficult to imagine how parties could negotiate the terms of rights after the injury had already occurred, when there had not been an initial assignment of the right. However, this does not mean that substantive rules are not subject to ex ante contract analysis. See *infra* note 170.

We can also see how an election would remove the “stickiness” of a party’s contentiousness, that is, cases where parties continue to litigate to gain incremental advantage, even though they generally agree on the factors of value and fully anticipate eventual settlement. The stickiness of contentiousness arises from a desire to uncover remaining unknowns, from the eternal hope for a better possibility. An election could eliminate this common barrier to resolution by signaling a party’s good faith belief in the superiority of his claimed right. If the parties generally agree on the factors of value, whether in mutual dialogue or in private reflection, and the differences are now minor, then the act of election and bonding could represent the final confirmation of this understanding and, given the increased stakes, prompt an expeditious settlement.

C. Judicial Administration

Any procedural innovation should be feasible and practical. A few comments on judicial administration and working details of the election scheme are needed at this juncture.

Simultaneous Disclosure. Fee shifting should apply only from the point of election, and all previously accrued fees should not be recoverable. This encourages, or at least does not discourage, some degree of information gathering and exploration of case merit. Simultaneous disclosure of election is required to ensure procedural symmetry. Serial disclosure produces unnecessary information asymmetry, resulting in unfairness, gamesmanship, and costly judicial intervention. Since election is not an advisable choice in many instances, parties may coordinate simultaneous disclosure by producing joint disclosures.⁸³ Lastly, simultaneous disclosure is a safeguard. A party can unilaterally disclose early. The benefit is the early accrual of fees subject to shifting, and the detriment is obviously the strategic drawback of giving the opposing party information and tactical options.

Irreversibility of Choice. Once an election is made, it should be irreversible. The bond of good faith is the assumption of a higher standard of proof. A bond is useless if it can be canceled at will or upon the occurrence of a callable event. In the same vein, an election scheme would be meaningless, and quite costly, if the parties could

⁸³ There are some game theoretic possibilities when a party rejects joint disclosure, but the strategies and outcomes are contextual and not subject to general observation. For example, without knowing the context, it is impossible to deduce how a party would or should construe the opposing party’s refusal to enter into a joint disclosure of nonelection. The possibilities are that the opposing party intends to elect, is unsure about election, is positioning for settlement leverage, is enjoying gamesmanship, or is simply incompetent.

petition to change their elections because subsequent developments prove the original election to have been unwise.⁸⁴

Factfinding. From the perspective of judicial administration, working with two standards of proof should not be an issue, as trial and appellate courts routinely do this. The application of two standards of proof requires particular findings of facts. A court can administer this task through either a special verdict form or a modified general verdict form.⁸⁵ The factfinder may be required to apply different standards of proof to different parties, claims, and defenses. This is a more complex task than the uniform application of the default standard. Nevertheless, there are many cases in which the factfinder is required to do this, for example, cases involving punitive damages.⁸⁶ With assistance from the court by way of clear jury instructions, the factfinder is capable of parsing parties, claims, and defenses under different standards of proof.

As usual, a factfinder should have wide latitude in deciding cases under an election scheme. Given its broad prerogative, it could view the plaintiff's failure to present a *prima facie* case under the default standard as a defendant's win by clear and convincing proof, that is, the defendant's heightened standard. But this is unlikely. In most cases, a plaintiff's failure to prove a case under the default standard, in and of itself, should not result in a conclusion that the defendant satisfied his burden of proving no liability under the heightened standard. Where there are several possibilities of past occurrence with different legal consequences attached to each theory of the case, the failure to prove a *prima facie* case under one party's theory does not necessarily mean that the negative proposition has been proven to the exclusion of other possibilities. For the defendant to win under a heightened standard, he should present affirmative evidence of no liability. This can be evidence disproving the plaintiff's theory of the case, coupled with the plaintiff's failure to carry her burden under the default standard, or evidence proving an affirmative defense under the heightened standard. Thus, the nature of the parties' obligations in the litigation system does not fundamentally change under the election scheme proposed here—the plaintiff must prove her case and dis-

⁸⁴ Of particular concern is satellite litigation over whether a party should be allowed to change his mind. Cf. George Cochran, *The Reality of "A Last Victim" and Abuse of the Sanctioning Power*, 37 LOY. L.A. L. REV. 691, 693 (2004) (noting that fee shifting under Rule 11 creates incentives for satellite litigation). Satellite litigation seeking to undo a bad choice is precluded if an election is irreversible as a matter of law.

⁸⁵ See FED. R. CIV. P. 49(a)–(b) (allowing special and general verdict forms).

⁸⁶ See *supra* notes 26–27 (providing examples of use of heightened standard in various contexts).

prove the defendant's defenses under the chosen standard, and the defendant has the opposite burden.

Multiclaime Suits. In a two-party, multiclaime litigation, there is no reason why each claim and defense should not be subject to per claim election rather than per party election. Election per claim is the better rule because the parties can more finely allocate the risks and returns of each claim and defense. This scheme promotes efficiency by incentivizing the settlement or dismissal of marginal claims and defenses. The start of a simple action can often devolve into a fusillade of charges and countercharges, many of which are marginal. The possibility of fee shifting on each claim and defense would deter shotgun litigation tactics. Obviously, the administration of a multiclaime action would be more complex. Individualized election makes the task of sorting facts, evidence, and evidentiary standards more difficult. Also, where two claims are tried and each party wins on different claims, the court must shift fees on a per claim basis. This raises the specter of litigation on attorney fees, increasing the cost of an action. But such litigation is common whenever there is fee shifting. Given the extent of the social cost from frivolous actions and unprosecuted meritorious actions,⁸⁷ an educated guess is that on balance, the benefit of deterrence of frivolity through fee shifting outweighs the added administrative cost.

Multiparty Suits. The complexity in a multiparty litigation may make an election scheme ultimately infeasible. There are myriad potential Gordian knots and contradictions. Consider a simple example: Plaintiff *A* sues defendants *B* and *C*, who were allegedly acting in concert.⁸⁸ *A* elects against *B* but not against *C* because she believes that she has a better claim against *B* in the event a factfinder finds that the defendants were not acting in concert. But *C* elects against *A*. At trial, *A* loses against *B*, even though the jury finds that *B* and *C* were acting in concert and that there was a preponderance of the evidence to support liability against both. It also finds that *C* failed to prove his defense under the heightened standard. The end result is that *A* loses as to *B*, despite the preponderance of the evidence against *B*, but wins against *C*. Assume that joint and several liability attaches to parties who injure a person while acting in concert. Can *C* seek contribution from *B*? Clearly, this situation involves the

⁸⁷ See *supra* Part I.B.

⁸⁸ See, e.g., *Keel v. Hainline*, 331 P.2d 397, 400-01 (Okla. 1958) (holding that defendants who aided or abetted harmful act are liable).

interplay of the election scheme with other areas of the law,⁸⁹ leading to complexity and uncertainty.

This simple hypothetical illustrates the many potentially thorny situations that can confront parties and courts in multiparty litigation. The complexity of multiparty, multiclaim litigation increases nonlinearly with additional claims and parties.⁹⁰ The administrative cost associated with resolving such complexity outweighs the benefits of private ordering. The potential to destabilize well-established rules of law in other areas is high, and the resulting uncertainty would introduce more cost into an already costly litigation system. When the pros and cons are weighed, the election scheme may well be inappropriate for multiparty litigation. With that said, it is plausible that given the social benefits, an election scheme could be applied in relatively simple multiparty litigation, though this category is easier to conceptualize than to realize. Such a decision should be made at the discretion of the court with appropriate input from the parties. We suspect, however, that a consensus among the court and parties would be difficult to reach, and as a practical matter, the application to multiparty suits would be difficult.

Timing of Election. The timing of the election is perhaps the most crucial administrative matter. There are two competing policies at work. On the one hand, parties should make their election with the most complete information possible. Discovery produces information and generally results in a diminishment of factual ambiguity.⁹¹ It is not unusual to see many attorneys and clients reluctant to settle early when there is significant factual ambiguity, choosing to settle only after some discovery is had. This policy consideration suggests that election should be made at the end of discovery but before any adjudication, including summary judgment.

On the other hand, all else being the same, the parties should minimize transaction costs. The completion of discovery is perhaps the most costly cash expenditure in litigation. If an election is made at the end of discovery, the scheme would have less influence on the goal of cost mitigation. A court must also be cognizant of the unintended consequence that, to bolster the possibility of proving their cases

⁸⁹ Even a two-party litigation could involve issues of res judicata and collateral estoppel if there is a potential party not involved in the litigation.

⁹⁰ Economists have noted that the transaction costs of negotiating an outcome increase as the number of participants increases. See Demsetz, *supra* note 30, at 354–55 (“Negotiating costs will be large because it is difficult for many persons to reach a mutually satisfactory agreement, especially when each hold-out has the right to work the land as fast as he pleases.”).

⁹¹ But see *infra* note 118 (noting that some scholars hold that egocentric interpretation of facts by parties will cause beliefs to diverge rather than converge).

under the heightened standard, parties electing to shift fees would be motivated to engage in more unwise discovery—unwise given the marginal cost. This problem is a lesser concern in frivolous and low value cases because the merits are clearer.

The balance of these competing policies suggests that an election should be made at some time *during* discovery. The discovery plan and the timing of election would be subject to the court's discretion. The court could bifurcate discovery. The first stage allows the parties to gather sufficient information to make an informed election, though certainly not with perfect information. Upon completion, the parties would make their election and then proceed to the second stage, which would focus on proving their case under the elections made.

Lastly, note the role of information. If there is significant information asymmetry, or symmetry with completeness of information, there is a distinct possibility that a party with a high degree of favorable information, such as the defendant in a frivolous suit or a plaintiff in a low value case, may choose to elect early in order to lock in fee shifting before expensive discovery is undertaken.

D. Conditions of Election

An election materially changes the prospects of a case, and thus the decision to elect is no small matter. The decision is a complex calculus. It is clear that the exercise of the option would *not* be standard fare in litigation strategy. In the vast majority of meritorious disputes, the parties would maintain the procedural status quo. Several factors would determine the likelihood of election: (1) the ratio of loss value to transaction costs, (2) a party's probability assessment, (3) several aspects of risk allocation, (4) a party's risk preference, (5) the level of information available, and (6) attorney incentives. These factors are discussed in order.

Ratio of Loss Value to Transaction Costs. The importance of this ratio is intuitively obvious. As the loss value increases, the transaction costs become less influential in the calculation of net value. The dispute becomes more sensitive to disagreements on the probability of success and expected value.⁹² All else being the same, a \$1 million case is more likely to go to trial than a \$100,000 case because a small degree of difference in views in the former case can still result in a large disputed amount that may exceed the cost of litigation. One is less inclined to elect a higher standard as the loss value increases

⁹² See Posner, *supra* note 35, at 419 n.29 ("There is empirical evidence that higher stakes do increase the likelihood of litigation.").

because the reduction in expected value associated with a higher standard of proof tends to offset the increase in potential fee recovery.

A numeric example illustrates the point. Assume that $L = 100$, $T = 10$, and $P = 0.6$. The ratio of loss value to transaction costs is 10x. Under the American rule, the plaintiff's net value is 50, calculated as: $V = PL - T = 0.6 \times 100 - 10$. Suppose the plaintiff elects, but the defendant does not; thus, the plaintiff arranges a one-way fee shifting scheme. The imposition of a higher standard of proof necessarily means that probability of success is diminished, though the quantum depends on the unique circumstances of each case. Assume, for example, that the probability decreases to $P = 0.5$ due to the assumption of the heightened standard. Under the payoff formula for a one-way fee shifting, the plaintiff's net value is now 45, calculated as: $V = P(L + T) - T = 0.5(100 + 10) - 10$. An election would result in a lower net payoff for the plaintiff.

Consider now a case in which the transaction costs are higher: $L = 100$ and $T = 25$. The ratio of loss value to transaction costs is 4x. Here, cost influences the party's decision and strategy more. If $P = 0.6$, the plaintiff's net value under the American rule is 35, calculated as: $V = PL - T = 0.6 \times 100 - 25$. Assume that the plaintiff elects and the probability decreases to $P = 0.5$ again. The net value of a one-way fee shift is now 37.5, calculated as: $P(L + T) - T = 0.5(100 + 25) - 25$. An election would result in a higher payoff.

These illustrations show that an election is more likely when the ratio of loss value to cost is lower, because the benefit of fee indemnity tends to offset the cost of diminished expected value and vice versa. Since smaller cases tend to have lower ratios of loss value to cost, they would generally enjoy a systematic benefit from the election scheme proposed (all else being the same).

Probability and Expected Value. A shift to a higher standard of proof always lowers the probability of prevailing, and thus the expected value. Imagine how the expected value of a civil action diminishes if the standard of proof changes to the criminal standard. A party would only elect if the reduction in the expected value, due to a decrease in the probability of liability, is exceeded by the expected gain from the recovery of fees.⁹³ Let us continue with the preceding

⁹³ We can represent this algebraically. For a plaintiff, an election to the heightened standard would result in the following: The value changes from the American rule, as represented by $V = PL - T$, to a one-way fee shifting scheme of $V' = P'(L + T) - T$, where V' is the new net value under the lower probability P' attributable to a heightened standard. A plaintiff would only elect if the new net value is greater than the old net value, such that $V' > V$. Substituting terms, we get the following inequalities: $P'(L + T) - T > PL - T$, or equivalently $P'T > L(P - P')$. This last inequality says that

example. It shows that upon election, the plaintiff's net value increases from 35 to 37.5 even as probability decreases from $P = 0.6$ to 0.5. If the probability decreased further to $P = 0.48$, the net value would be 35, calculated as: $P(L + T) - T = 0.48(100 + 25) - 25$. The net value under the one-way fee shifting would be exactly equal to the status quo under the American rule, suggesting that $P = 0.48$ is the indifference point between the status quo and election (assuming risk neutrality). Any further reduction in probability below this threshold would result in a diminishment of net value. Thus, an electing party must carefully consider the relationship between fee recovery and reduction in expected value due to a higher standard of proof.

Aspects of Risk Allocation. By changing the standard of proof and the attorney fee rules, an election fundamentally changes the risk profile of a case. It not only lowers probability but can also increase the riskiness of a case outcome. Unless a party has brought an absolutely frivolous lawsuit (certainty of loss) or has bribed the judge and jury (certainty of win), there is always a range within which one's probability estimate may deviate from expectation. Two sets of facts and laws can produce the same probability assessment from the parties, but parties may assign different confidence levels to that assessment based on their perception of the completeness or ambiguity of information.⁹⁴ Thus, for any given assessment there is a margin of error.

The case outcome may be more sensitive to variance (potential deviations from one's expected value) at certain levels of probability than at others. The intuition is that high probability cases are less sensitive to the same level of variance than close cases.⁹⁵ In the former situation, even if one is uncertain of the precise location of the decision standard, there is a significant margin of error such that relatively small miscalculations do not affect the overall result. In the latter, a small deviation from expectation can change a close win to a close loss. The case is more sensitive to variations in expectation.⁹⁶

There is another aspect to the risk analysis. Risk can be measured as the variance of net values under different fee rules. Under a

election is expected when the anticipated gain in transaction costs $P'T$ is greater than the reduction in expected value $L(P - P')$.

⁹⁴ See Rhee, *supra* note 38, at 653–61 (providing example and analysis of how weight of, variance of, and confidence in one's probability assessment can influence party's choice and strategy).

⁹⁵ See Priest & Klein, *supra* note 61, at 16–19 (providing model showing that probability and variance are important to each party's expectation of outcome and case value).

⁹⁶ See Rhee, *supra* note 38, at 635 n.67 (providing numeric example of Priest and Klein's concept).

one-way fee shifting scheme, where only one party elects, the range of returns is skewed in favor of the electing party. This change in risk is positive and thus always beneficial to the electing party. As discussed, however, a shift from the American to the English rule always causes the range of payoffs and payouts to be wider. Thus, a bilateral election results in greater variance of returns.

In sum, fee shifting increases the risk in several important respects. There is variance of returns, which is not a good thing for a risk-averse party, unless the variance increases only on the upside, as would be the case in one-way fee shifting. But more problematic is the influence of risk on one's confidence. A party must be satisfied that a substantial reduction in probability would still yield a win, and as the probability decreases, a party would be more sensitive to potential deviations from expected value. This suggests that a party would elect only when the level of information is high, providing greater confidence in the estimates of probability and risk.

Risk Preferences. A risk-neutral party conducts a single variable calculation based solely on how a decision affects the probabilistic outcome. But in a scheme in which the standard of proof and attorney fee shifting is optional, there is not a single expected value. As discussed, the essential financial calculus is whether the reduction in probability is offset by the potential gain of fees in the event of a win.⁹⁷ This is a more difficult assessment because it requires a second probability analysis and a relative comparison of the two net values.

As for non-risk-neutral parties, there are several additional considerations. Although a one-way fee shifting skews the range of outcomes upward in one's favor, the probability of success must decrease because of the higher standard of proof. Under the American rule, a risk-averse party may choose a lower net return with a higher probability of winning than a higher net return under a one-way rule. In other words, the risk-averse party may not equate the same probabilistic expectations of net values under different fee rules, but in fact may value a higher probability of success (expected value) more than the total expected return (net value). The consideration for a risk-seeking party is the opposite: A risk-seeking party would sacrifice a quantum of return to opt for the more risky choice. Thus, for both the risk-averse and the risk-seeking parties, there is a point at which the risk and net value are at equilibrium and they are indifferent as to the valuational consequences of election.

⁹⁷ See *supra* notes 92–96 and accompanying text.

Level of Information. The most important problem in dispute resolution concerns information acquisition and asymmetry.⁹⁸ Information, the basis of valuation, is costly to acquire. It would be foolish for any party to attempt a valuation based on limited knowledge, and one suspects that the loud, if not cacophonous, complaints of irrational or inefficient behavior in the litigation system are based largely on ex post assessments of outcomes that are the product of necessarily imperfect ex ante predictions and decisions occurring in a state of uncertainty about future outcomes.⁹⁹ The process of valuation is the incorporation of all available information. In a world of uncertainty, a sufficient level of information produces confidence in one's valuation. Only when there is proper pricing can parties engage in a more precise dialogue and effectively resolve a dispute.¹⁰⁰

In this cloud of uncertainty, the effect of self-interest diminishes the quality of information flow. It is certain that parties are never neutral and that they always want to win. A party will disclose information only if it is mandated or serves one's interest, or at least this is the strong presumption. Typically some degree of information asymmetry exists. The parties are uncertain as to the level of information that would constitute full disclosure of all knowable facts (for practical purposes, a state of perfect information).

Many years ago, Frank Knight distinguished the concepts of risk and uncertainty: risk connoting a future state of known distribution, and uncertainty connoting a future state of unknown distribution or imperfect information.¹⁰¹ A similar concept is applicable in the case of information asymmetry. Even when there has been full disclosure, there may be trepidation as to the level of uncertainty. A party with less information can never be certain that the opposition has disclosed all information. Imagine how simple negotiation would be if parties

⁹⁸ Some of the most important economic concepts of the past thirty years have been information dispersion and uncertainty. Interview with Prof. Kenneth Arrow, Stanford University, in *Three Nobel Laureates on the State of Economics*, CHALLENGE, Jan.-Feb. 2000, at 6, 13, 19-20.

⁹⁹ See *supra* note 61. Too often, scholars, policymakers, and judges apply a hindsight bias to difficult decisions when they have no stake in the value, cost, and risk. "Once the outcomes are observed, it usually is easy to say what would have been the best decision." HARRINGTON & NIEHAUS, *supra* note 63, at 35. Likewise, in a casino or the capital markets, it is easy to assess strategies of value, cost, and risk when there is nothing at stake other than one's own mental accounting of what one would have done.

¹⁰⁰ See Rhee, *supra* note 38, at 653-61 (discussing interplay of parties' confidence regarding projections and settlement decisions).

¹⁰¹ KNIGHT, *supra* note 63, at 233-34.

could somehow read each other's minds.¹⁰² But of course, uncertainty and strategic behavior are facts. Nor can we attribute to disputants the unnatural quality of omniscience for the purpose of normative prescription.¹⁰³ Thus, the road to achieving a level of information parity must be paved with mandatory disclosure rules and the attendant expenditure of transaction costs.

At each point in the process of litigation, the risk perceived is a function of the available information, the perceived unknowns and unknowables, and the expected amount of future disclosures to be made through the litigation process. When a lawsuit is filed, the level of known information is typically low. Under modern pleading doctrine, a party need not engage in extensive discovery to file a pleading. Unless specifically mandated,¹⁰⁴ the doctrine does not require particularized averments.¹⁰⁵ It only requires a "short and plain" statement of the claims and defenses,¹⁰⁶ which can be contradictory without violating ethical obligations.¹⁰⁷ Although a party is subject to the good faith obligation of Rule 11, he is not judged on ex post accuracy but on ex ante reasonableness.¹⁰⁸ As litigation moves from the pleading to discovery stage, the level of information generally increases.¹⁰⁹

Because the difference in the level of information between pleading and post-discovery is substantial, the risk is generally much higher in the early stages of litigation.¹¹⁰ Only upon final judgment is there certainty. Before this, probability is based on the information available. The unknown and the awareness of it have no effect on

¹⁰² This condition would represent a state of perfect information, that is, a state of perfect knowledge by all exchangers of products. *Id.* at 86. "If intercommunication is actually perfect, exchanges can take place at only one price." *Id.* at 82.

¹⁰³ "It can never be assumed that parties should try to accurately predict the decision of the deliberative body because this assumes a level of predictive power beyond the credible allowance of a rational person." Rhee, *supra* note 38, at 664; *see also supra* note 61 (discussing rational expectation model).

¹⁰⁴ FED. R. CIV. P. 9(b), 23.1(b)(3).

¹⁰⁵ FED. R. CIV. P. 8(a), (b)(3), (b)(5).

¹⁰⁶ FED. R. CIV. P. 8(a)(2), (b)(1)(A).

¹⁰⁷ FED. R. CIV. P. 8(d)(3).

¹⁰⁸ FED. R. CIV. P. 11(b). *See generally* Charles M. Yablon, *The Good, the Bad, and the Frivolous Case: An Essay on Probability and Rule 11*, 44 UCLA L. REV. 65 (1996) (arguing that frivolous cases under Rule 11 are not necessarily meritless, but instead are generally believed by plaintiffs ex ante to be reasonable "long shots" with low, but not zero, probability of success).

¹⁰⁹ *See Rhee, supra* note 38, at 676–78 (showing effects of progress of litigation on parties' level of uncertainty).

¹¹⁰ *See id.* at 676 & n.246, 677 (showing that with limiting assumptions, valualational impact of new information is subject to declining marginal utility curve).

probability since it cannot be assigned a probabilistic value.¹¹¹ As the litigation progresses, we expect a random walk of probability as previously unknown information is discovered.¹¹² Thus, probability is always subject to variance, whose range is determined by each party's perception of the degree of the unknown and unknowable information.

Consider, then, two different sets of circumstances in which the probabilities are the same, except that the level of uncertainty differs, as in pre- and post-discovery states of case assessment. In the pre-discovery scenario, the probability is high, but the uncertainty is high as well. As a result, the high expectation of positive outcome is tempered by the reality that it is subject to a greater range of variance from expectation. At the lower range of this variance, the party could lose because the case falls short of the court's decision standard demarking the boundary between liability and no liability, and this loss is within the range of expectation when one considers the large variance due to uncertainty. In the post-discovery scenario, one would have more confidence, and the spectrum of variance would be tighter with the same measure of high probability as before discovery. Here, the lower end of the spectrum of variance could exceed the court's decision standard if the variance is low enough such that the high probability is coupled with greater confidence in this estimate.

The point is that a party would be ill advised to elect a higher standard of proof under the condition of low information. Where the level of information is high, such as at the end of discovery, rational parties have greater confidence in their views. If the parties view the case similarly and the risks have been sufficiently mitigated through some discovery, the parties most likely will end up settling the case. Barring the utility of gambling, a high level of information is an essential condition for election.¹¹³ This implies four aspects of election pertaining to the relationship between information and risk.

First, the timing of an election is important. An election is more likely when there is a high degree of disclosed information. A probability assessment in the early stages of litigation, even if it suggests that the English rule is better at that point in the litigation pro-

¹¹¹ For example, the parties anticipate deposing a witness whose proposed testimony is unknown to the parties. This unknown factor cannot affect probability because the parties cannot assign it a value. *Id.* at 654.

¹¹² *Id.* at 664-66.

¹¹³ See Priest & Klein, *supra* note 61, at 17 ("In litigation, as in gambling, agreement over the outcome leads parties to drop out.").

cess, would be subject to a low degree of confidence.¹¹⁴ Election under such a condition would generally be speculation rather than an informed choice.

Second, if information asymmetry exists, the party with better information is more likely to elect, particularly when he knows that the opposition is substantially disadvantaged. From the point of view of the electing party, she would want the option to elect, but not have an election against her, particularly if the opponent's election is based on sound case assessment and information. Therefore, as a general proposition, unilateral fee shifting based on asymmetric information advantage is always better than bilateral fee shifting based on completeness and symmetry of information between the parties. An election puts significant pressure on the other party, who lacks the same degree of information, which may result in that party decreasing its settlement value.

Third, election may serve an important signaling function that conveys essential information about the parties' assessments, confidence level, and evidence. Because an election is bonded, the level of noise in this signal is low and the information quality is high.

Lastly, a high probability, low variance case is more likely to see election. A good example is a case of negligence for the amputation of a wrong limb or organ. Liability in these cases is clear, and if there is a dispute, it is usually over damages. Although the parties usually settle these cases (because they would generally agree on the issue of liability and because the risk is low), election provides a substantial benefit. The compensation to the injured party would then be whole, as attorney fees do not subtract from the recovery.

Attorney Incentives. The interest of the attorney must be considered in any procedural reform. Litigation is a team production between attorney and client. In matters of strategy, the attorney plays the predominant role. Although the attorney is bound by the rules of ethics to consider the best interest of the client, she is nevertheless an agent in a principal-agent relationship. Economists tell us that there is no perfect agent.¹¹⁵ The attorney's interests are simple: She wants to be paid and does not want to lose the case. The interest in the former is strictly her interest, while the interest in the latter is mixed. Obviously, winning a case serves the client, but it also serves the attorney's interest. These interests are intimately related. Winning enhances the

¹¹⁴ See Rhee, *supra* note 33, at 240 ("Confidence is intimately linked to the perception of risk."); Rhee, *supra* note 38, at 653-55 (explaining that confidence and risk are correlated in that low risk results in higher confidence in one's expectation).

¹¹⁵ See JENSEN, *supra* note 11, at 5 ("[P]eople will not act in the interest of others (their principals or partners) to the exclusion of their own preferences.").

attorney's reputation, increases the possibility of repeat business from the client, and in a contingent fee arrangement, assures fees.

The interests of the attorney and client can diverge. The client seeks maximum recovery at the lowest cost. This may result in a preference for election, which lowers the probability of success but increases the net value. But the attorney will generally not prefer this choice. Under a contingent fee agreement, the likelihood of recovery is reduced, but the judgment for damages from which the fee is taken remains the same. This reduces the attorney's economic incentive. Under an hourly fee arrangement, the fees presumably do not change; only the paying party changes. If there is no difference in the credit quality of the paying party, the attorney has no economic incentive to prefer election, and she faces the prospects of an increased chance of losing the case. We expect that the disincentive toward election would flow through to the attorney's advice to the client, though it would be improper to suggest that agent self-interest always or even frequently overrides the agent's fidelity and duty to her client's interests.¹¹⁶

Notwithstanding the agent's conflict, there are several circumstances in which an attorney would have an economic incentive to elect the heightened standard. First, in the case of frivolous action, the defense counsel may not be so concerned with the prospect of losing, even under a higher standard of proof, and a successful defense of the client's case under an even higher standard of proof may burnish the attorney's reputation, as well as provide an opportunity for repeat business. Second, an attorney may have an economic incentive for election if an election creates the possibility of a new business opportunity. That new business opportunity is the low value case. Consider, for example, a hypothetical case in which a plaintiff has clearly suffered an injury worth \$25,000 but would have to expend \$50,000 in legal fees to bring the action in court. In this case, the possibility of fees, though risky, is better than the certainty of no fees. These cases would ordinarily not be prosecuted absent a class action or consolidation. The possibility of fee shifting opens up a class of traditionally excluded or marginalized cases.¹¹⁷ Here, the interests of the attorney and client align more closely, as each seeks the pursuit of a common opportunity.

¹¹⁶ See *id.* (noting that rational self-interest is frequently misinterpreted to mean pathological selfishness with no leeway in preference for well-being of others).

¹¹⁷ See *infra* Part III.B (providing more analysis of impact of election scheme on high probability, low value suits).

* * *

It is helpful to summarize the conditions of election. Clearly, an election fundamentally changes the essential attributes of a case. Most cases will not result in election. Parties will be wary of exercising the option, particularly where the attorney incentives militate against an election. An election necessarily reduces the expected value of a case (because it reduces the probability of winning), and there is an economic benefit only if the benefit of fee indemnity outweighs the diminishment of expected value. Also, an election reconfigures the risk profile of a case. Parties will elect only when there is high confidence in their assessment of the case, which is correlated to a high level of probability and information. This set of unique conditions is met in two types of cases—low value actions and frivolous suits.

Since a high probability of success is a natural condition of election, we can fairly predict that the English rule will be seldom invoked through bilateral choice once discovery has commenced. For this to occur, both parties must believe the probability of success is high and must have a high degree of information such that the perceived risk is low. Although rational parties can disagree about predicted outcomes and interpretations of facts and laws, and such disagreements can easily lead to trial without normative condemnation that the trial is the result of some “error,”¹¹⁸ it is unlikely that disagreement on probability and information is so stark as to result in bilateral election.

The most likely situation is no election at all in the vast majority of cases, resulting in the status quo of the American rule and the default standard of proof. An election by one party is likely in a small but socially significant class of cases, resulting in potential one-way fee shifting. Lastly, bilateral election, resulting in the potential application of the English rule if one party meets the heightened standard of proof, would almost never occur so long as the timing and structure of the election scheme are adequately planned, as discussed above.¹¹⁹

¹¹⁸ See George Loewenstein & Don A. Moore, *When Ignorance Is Bliss: Information Exchange and Inefficiency in Bargaining*, 33 J. LEGAL STUD. 37, 37 (2004) (“[S]hared information, if open to multiple interpretations, is likely to be interpreted egocentrically by the disputants, which can cause beliefs to diverge rather than converge.”).

¹¹⁹ See *supra* Part II.C.

III

SOME IMPLICATIONS OF A BONDED UNILATERAL FEE-SHIFTING REGIME

A. *Frivolous Lawsuits*

The policy implications of frivolous lawsuits are broad.¹²⁰ There is a common perception, right or wrong, that the justice system entertains too many frivolous actions.¹²¹ The responses to frivolous actions largely have been consistent, differing only in the details. Aside from the occasional calls for greater professional responsibility and ethics,¹²² most scholarly analyses have focused on fee shifting when a case is determined by the court to be frivolous.¹²³

We can refine the problem further. To the extent frivolous actions are a systemic problem, they are most problematic in the early phases of litigation. They rarely go to trial, and the problem is the occurrence of needless cost in the beginning, followed by an extortionate settlement or dismissal later. Cases that are determined to be meritless only after discovery are not problematic. We should not make an ex post judgment on an ex ante decision, particularly in a system of notice pleading, for any reason other than to judge whether the decision to file suit was reasonable.¹²⁴ In this light, the proposition

¹²⁰ The problem of deterring frivolous lawsuits is much discussed in the literature. See generally Robert G. Bone, *Modeling Frivolous Suits*, 145 U. PA. L. REV. 519 (1997) (proposing model of sanctioning frivolous litigation); A. Mitchell Polinsky & Daniel L. Rubinfeld, *Sanctioning Frivolous Suits: An Economic Analysis*, 82 GEO. L.J. 397 (1993) (same).

¹²¹ See Valerie P. Hans & William S. Lofquist, *Jurors' Judgments of Business Liability in Tort Cases: Implications for the Litigation Explosion Debate*, 26 LAW & SOC'Y REV. 85, 96 (1992) (noting that research revealed "the widespread impression among jurors that the civil litigation system is overburdened by claimants seeking awards in meritless cases"). But see Deborah L. Rhode, *Frivolous Litigation and Civil Justice Reform: Miscasting the Problem, Recasting the Solution*, 54 DUKE L.J. 447, 459 (2004) ("Although excessive litigation is the pathology dominating public discussion and policy agendas, systematic research reveals that the more serious problems are undercompensation of victims and overcompensation of lawyers.").

¹²² See generally Neal H. Klausner, Note, *The Dynamics of Rule 11: Preventing Frivolous Litigation by Demanding Professional Responsibility*, 61 N.Y.U. L. REV. 300 (1986) (arguing that greater professional standards should mitigate frivolity).

¹²³ See Bebachuk & Chang, *supra* note 57, at 371-72 (proposing that courts should have power to shift fees based on adjudicated "margin of victory"); Nathalie D. Martin, *Fee Shifting in Bankruptcy: Deterring Frivolous, Fraud-Based Objections to Discharge*, 76 N.C. L. REV. 97, 103 (1997) (proposing that bankruptcy courts should have power to impose fees for improper objection to discharge).

¹²⁴ Cf. FED. R. CIV. P. 11(b) (stating that by presenting pleading to court, attorney or unrepresented party is certifying that "the allegations and other factual contentions have evidentiary support or, if specifically so identified, are likely to have evidentiary support after a reasonable opportunity for further investigation or discovery" (emphasis added)).

that the court should adjudicate the issue of frivolity assumes that it is the most competent decisionmaker to conduct this inquiry.

Stated differently, this view assumes that the court should determine the issue of frivolity because it is either neutral or best informed among all participants. The second explanation is simply wrong, as the court, being the most removed from the activities giving rise to the dispute, is the least informed and last to know. The only justification then is that the court is a neutral arbiter, a rather uninspiring reason given its inherent lack of knowledge and the cost associated with achieving some degree of information parity with the parties. Indeed, this reason makes clear that the function of a court is somewhat fungible, as demonstrated by the prominence of arbitration and other alternative dispute resolution mechanisms as substitutes for public adjudication.

In truth, the parties are in the best position to recognize a frivolous suit. This is obvious since frivolity implies knowledge; a truly frivolous case is brought by a plaintiff who *knows* she has no legitimate chance of success on the merits. The problem, then, is one of information conveyance and its cost. Strategic behavior and self-interest undermine an actual, honest dialogue, which is needed for an efficient result. Any settlement dialogue has much noise, which becomes even more extreme when a plaintiff has a pecuniary interest in maintaining a frivolous suit and making a credible threat. This presents a policy question: How can the law elicit honest assessments from the parties regarding case merit? Such an assessment would allow the parties to settle their dispute at a cost that would reflect an honest hypothetical dialogue between them, were the effect of self-interest somehow nullified.

Thus far, the assumption in the literature and in the rules of civil procedure has been that the power to deter frivolity resides with courts. Judicially imposed fee shifting certainly could deter frivolous litigation, but it does so at the cost of actual adjudication of the issue or claim. The result is typically satellite litigation used to determine whether the original action was indeed frivolous.¹²⁵ Indeed, some adjudication of the issue of frivolity is necessary to establish precedent and to give guidance on the location of the decision standard. Also, such fee shifting may cast too broad a net, capturing cases determined *ex post* to be meritless, even though frivolity was not the plaintiff's ex

¹²⁵ See, e.g., *Chambers v. NASCO, Inc.*, 501 U.S. 32, 33 (1991) (expressing concern over "extensive and needless satellite litigation" in Rule 11 inquiry); *Cooter & Gell v. Hartmarx Corp.*, 496 U.S. 384, 408 (1990) ("Moreover, including appellate attorney's fees in a Rule 11 sanction might have the undesirable effect of encouraging additional satellite litigation.").

ante intent. Thus, adjudication of frivolity is a blunt instrument that undermines the policy of cost mitigation.

Assume now that a class of litigants, deliberating behind a Rawlsian veil of ignorance, must arrange for the elimination of frivolous claims. They would have a choice between two normative ends: either shift costs to the adjudged frivolous party or minimize total cost expenditure. These goals are not coterminous. Once a suit is filed, a party cannot judge the merits of the other's action, and so the parties must turn to an independent third party (the court) to determine the issue of fee shifting at a potentially high cost of resolution. Since the parties do not know the future assignment of the cost, they would agree that the better goal is the deterrence and termination of frivolous cases at total minimum cost. This suggests that early in litigation, which is the critical timeframe, information acquisition and conveyance to the court should be limited. In other words, the parties would agree—*do not litigate*.¹²⁶ The election scheme proposed here could help parties to achieve this goal by incentivizing earlier settlements.

The proposal here is consistent with the goal of total cost mitigation. If a case is frivolous and both parties understand it to be, they should come to an implied agreement on this point. The defendant would want fee shifting. But he should not be able to do this absent an assurance that the unilateral imposition of fee shifting is made in a good faith effort to avoid further harm. The cost of this bonding is the imposition of a higher standard of proof, which should be the most efficient price of bonding if the action is indeed meritless. Conversely, the unilateral imposition of potential fee shifting requires that the plaintiff also provide a bond of her good faith in an implied bargain with the defendant. She can do this by proving her case under the default standard or by defeating the defendant's attempt to prove his case under the heightened standard. This is impossible if her case is in fact frivolous. Thus, if the parties are allowed to privately order the procedural rules to better reflect the value, risk, and cost profile of a case, the system of public adjudication can help the parties reach a more efficient settlement.¹²⁷

¹²⁶ This is an outcome that resonates with most scholars, judges, and policymakers. See Gross & Syverud, *supra* note 39, at 3 (noting American judicial system's premise that "[t]rial is a disease, not generally fatal, but serious enough to be avoided at any reasonable cost").

¹²⁷ The scheme advanced here is not effective if the penalty has no deterrent effect. A judgment-proof frivolous plaintiff is problematic in the abstract. In reality, however, it is an open empirical question as to how much frivolous litigation actually is brought by such plaintiffs. The judgment-proof plaintiff conjures up the image of assetless, opportunistic, and perhaps desperate people whose subsistence is to feed off the litigation system and innocent (typically corporate) victims. For this author, this stereotype, perhaps engen-

B. Low Value Cases and Class Actions

The low value case is essentially an inverse variant of a frivolous suit with respect to probability. A dynamic similar to that described above works for low value cases. These cases also present other considerations. If a plaintiff with a low value claim does not want to risk significant cost, she must find a means by which to spread the risks and costs. The class action was designed for this type of situation,¹²⁸ but it is not a readily available option. An attorney must be willing to undertake such a large endeavor, and an action must meet substantial and onerous legal requirements.¹²⁹ One-off low value disputes cannot qualify for class action treatment. If there is not an existing action, there are search and startup costs of finding an attorney and commencing an action. Most aggrieved people will not undertake this endeavor. Thus, the potential plaintiff of a low value dispute is generally subject to the unlikely prospect that a class action exists and that she has been identified as a class member.

If a class action exists, there are still other considerations. Many class actions, like many individual actions, succumb to the pricing pressures of settlement.¹³⁰ In individual cases, the plaintiff can be actively involved in the settlement discussion, providing direct input into the issues of compensation and justice, though we recognize that the attorney has a substantial, if not prominent, voice. But the class action is administered by the class counsel, the defendant, and the court, generally to the exclusion of absent class members. It is often reduced to a basic calculation of compensation and fees. On fees, commentators have noted that plaintiffs' attorneys generally get too much.¹³¹ In any attorney-client relationship, there is the problem of

dered by ideology and politics, seems cartoonish and does not rise to the quality of a general observation of the litigation system. Ordinary people do not like litigation; they have some assets and garnishable wages; and they are risk averse. These are the realistic aspects of our world, and reality disposes of the vast majority of the small minority of cases in which the cartoon in fact does come to life.

¹²⁸ See *Amchem Prods., Inc. v. Windsor*, 521 U.S. 591, 617 (1997) ("The policy at the very core of the class action mechanism is to overcome the problem that small recoveries do not provide the incentive for any individual to bring a solo action prosecuting his or her rights." (quoting *Mace v. Van Ru Credit Corp.*, 109 F.3d 338, 344 (7th Cir. 1997))).

¹²⁹ See FED. R. CIV. P. 23 (outlining numerous legal hurdles to clear in order to certify class actions).

¹³⁰ See, e.g., *In re Rhone-Poulenc Rorer Inc.*, 51 F.3d 1293, 1298–1300 (7th Cir. 1995) (Posner, J.) (analyzing pricing pressures in class actions); Rhee, *supra* note 41, at 150–54 (same). See generally Richard A. Nagareda, *Aggregation and Its Discontents: Class Settlement Pressure, Class-Wide Arbitration, and CAFA*, 106 COLUM. L. REV. 1872 (2006) (discussing class settlement pressure).

¹³¹ See, e.g., Rhode, *supra* note 121, at 459 (noting problem of lawyer overcompensation). More generally, contingent fees have been criticized. See generally Bruce L. Hay, *The Theory of Fee Regulation in Class Action Settlements*, 46 AM. U. L. REV. 1429 (1997)

agency cost.¹³² Since the relational distance between class counsel and clients is even wider in the class action context, the problem of agency cost must be significantly greater.

Procedural optionality would provide an alternative to the traditional class action. Importantly, the plaintiff's prospect of compensation would no longer be substantially linked to the fortuity of whether a class action exists. The search and start-up costs associated with initiating a class action would be eliminated for plaintiffs who were willing to go it alone and elect to shift fees to the defendant. The plaintiff no longer would have the legal challenge of satisfying onerous class certification requirements.¹³³ She would control the resolution of her dispute (settlement or litigation) and decide the price of resolution, which may reflect pecuniary and nonpecuniary considerations. Lastly, agency cost would be reduced because there would be a direct attorney-client relationship.

That procedural optionality would provide an alternative to the class action may also influence the economics of class actions themselves. In a class notice, plaintiffs are advised that they can opt out of the class and pursue their own action. For most plaintiffs, this option is economically infeasible,¹³⁴ hence the existence of the class action to take advantage of economies of scale in the legal representation of numerous small claims. But suppose that an economically feasible alternative exists. Some plaintiffs may choose an individual action. Moreover, attorneys would take these individual claims if they had a reasonable chance of fee recovery. Thus, the election scheme creates a new business opportunity for this group of attorneys. If the merit of a case is strong enough, procedural optionality increases the supply of cases, increases the demand for non-class attorneys, and decreases the demand for class actions. As discussed, we want *more* litigation in the litigation category of low value cases. Competition in the prosecution

(analyzing principal-agent problem in class action settlements); Task Force on Contingent Fees, Tort Trial and Insurance Practice Section of the American Bar Association, *Report on Contingent Fees in Class Action Litigation*, 25 REV. LITIG. 459 (2006).

¹³² See also JENSEN, *supra* note 11, at 85–87 (discussing economic theory of agency cost); cf. Bruce L. Hay, *Contingent Fees and Agency Costs*, 25 J. LEGAL STUD. 503, 504–05 (1996) (applying agent-principal model to analyze proper fee levels).

¹³³ See, e.g., *Ortiz v. Fibreboard Corp.*, 527 U.S. 815, 864 (1999) (denying class certification based on failure to satisfy Rule 23(b)(1)(B) under theory of limited funds); *Amchem*, 521 U.S. at 628 (holding that requirement of commonality of issues of fact and law and adequacy of representation were not met in settlement-only class certification).

¹³⁴ See, e.g., Theodore Eisenberg & Geoffrey Miller, *The Role of Opt-Outs and Objectors in Class Action Litigation: Theoretical and Empirical Issues*, 57 VAND. L. REV. 1529, 1532 (2004) (presenting empirical findings showing that “on average, less than 1 percent of class members opt-out”).

of an action is good and would tend to reduce attorney fees and agency cost.

The suggestion here is not that a scheme of optional fee and proof rules would provide an alternative procedural outlet similar in scale to the class action. The benefits of procedural optionality proposed here work mainly at the margin of this class of cases. They are significant only if the case profile is suitable for election (that is, clearly meritorious), which limits many actions. But when a case satisfies the conditions of election, an individually pursued action may reduce the agency costs present in class actions, and thereby better serve the interests of the plaintiff and society.

C. *Tort Law as a Case Study*

Because private ordering of procedural rules can result in more efficient dispute resolution, procedural optionality can advance the policies and goals underlying substantive laws, as well as increase social efficiency. Tort law provides a good vehicle by which to test this proposition.

To begin, most meritorious tort cases will not be suitable for procedural election. Consider the domain of accident law. It is seldom the case that negligence is clear. Negligence is defined as the failure to exercise reasonable care under the circumstances.¹³⁵ Reasonableness is a highly malleable standard: “[T]he standard is known and thus not arbitrary, but its application is uncertain and thus unpredictable to a degree.”¹³⁶ The inability to precisely locate the decision standard that will be applied by the factfinder creates significant uncertainty.¹³⁷ When liability is a close question, election would be unwise, and only the foolish client or incompetent attorney would pursue it. Accordingly, the influence of an election scheme would operate at the margin of tort law—cases where the actions unambiguously should be dismissed or where liability is clear.

Procedural election would limit frivolous tort suits for the reasons discussed above¹³⁸ and would also better serve the broad policy goals of tort law in cases with a high probability of liability.¹³⁹ Although the

¹³⁵ See generally RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL HARMS § 3 (Proposed Final Draft No. 1, 2005) (providing definition of negligence).

¹³⁶ Rhee, *supra* note 41, at 172.

¹³⁷ See *id.* at 171–72 (noting that negligence standards, for example, are “difficult to fix or predict,” leading to uncertainty).

¹³⁸ See *supra* Part III.A.

¹³⁹ It should be noted that in special circumstances, when courts hold a certain behavior negligent as a matter of law, even negligence is said to be clear and without controversy. See, e.g., *Martin v. Herzog*, 126 N.E. 814, 815 (N.Y. 1920) (“We think the unexcused omission of the statutory signals is more than some evidence of negligence. It is negligence in

goals of corrective justice, compensation, and deterrence compete to be the theoretical center of gravity of tort law,¹⁴⁰ they are not mutually exclusive. Rather, these goals are mutually reinforcing on a basic level.¹⁴¹ Compensation tends to correct a wrong, and absent a legislative overhaul of tort law through some form of regulatory sanction or universal compensation scheme,¹⁴² compensation is also the primary means by which to deter inefficient activities. Where a case of negligence is clear, an election scheme could better serve the purposes of compensation and deterrence in a manner that is consistent with the principles of tort law.

The application of the American rule ensures that tort victims receive only partial compensation.¹⁴³ Since settlement is the primary means of resolving a tort dispute in practice,¹⁴⁴ compensation and deterrence are largely a function of some degree of private ordering in the shadow of the American rule, which the parties cannot change.¹⁴⁵ In the typical tort case, there is asymmetric risk: A corporate defendant is less risk averse than an injured plaintiff, it can diversify away the risk of a common case by holding a portfolio of similar lawsuits, and it has greater financial resources such that the opportunity cost of the stake at issue is lower than the plaintiff's.¹⁴⁶ This asymmetry of risk produces different valuations between plaintiffs and defendants,

itself."); *Theisen v. Milwaukee Auto. Mut. Ins. Co.*, 118 N.W.2d 140, 144 (Wis. 1963) (holding that falling asleep while driving is negligence per se).

¹⁴⁰ See generally John C.P. Goldberg, *Twentieth-Century Tort Theory*, 91 GEO. L.J. 513 (2003) (discussing various theories of tort law).

¹⁴¹ Cf. Gary T. Schwartz, *Mixed Theories of Tort Law: Affirming Both Deterrence and Corrective Justice*, 75 TEX. L. REV. 1801, 1834 (1997) ("As tort objectives, then, corrective justice and deterrence can be recognized as collaborators rather than competitors.").

¹⁴² See generally Rhee, *supra* note 41, at 170–81 (arguing that negligence persists because it promotes private resolutions of tort disputes); G. Edward White, *The Unexpected Persistence of Negligence, 1980–2000*, 54 VAND. L. REV. 1337, 1341, 1344 (2001) (observing that negligence has not been replaced with more public scheme of enterprise liability).

¹⁴³ See Keith N. Hylton, *Litigation Costs and the Economic Theory of Tort Law*, 46 U. MIAMI L. REV. 111, 114 (1991) ("A tort victim's cost of litigating consumes roughly thirty percent of the average damage award."); see also Rhee, *supra* note 41, at 164 ("Deductions from compensation are a structural feature of the tort system. The most obvious factor is the American rule of attorney fees. But a less obvious structural feature . . . is the risk-adjusted discount in settlement value." (footnote omitted)).

¹⁴⁴ It has been argued that the decision to settle versus litigate in the tort context does not affect the goal of efficient deterrence. See POSNER, *supra* note 32, at 48 (arguing that even though most cases settle, this does not affect efficiency of tort law because parties have incentive to collect information and reach "a reasonable settlement"). But see Rhee, *supra* note 41, at 168 ("Tort law is not efficient, as efficiency has been defined by the economic model, because courts are largely irrelevant in the instrumental function of determining value.").

¹⁴⁵ Rhee, *supra* note 41, at 127.

¹⁴⁶ *Id.* at 158–60.

even if they may view the facts and laws similarly.¹⁴⁷ This means that settlement valuations are typically discounted from the risk-neutral valuations.¹⁴⁸ Under this view, the plaintiff faces a deficiency in compensation on two fronts: self-financed attorney fees and valuation discount to the defendant.

This compensation scheme is firmly established in American tort law, and it would be a radical departure to suggest that fee shifting should be the general rule in tort law. Rather, an election scheme would create a sensible exception that works at the margin of tort law. Although many cases of negligence are ambiguous, there are some acts where negligence is clear, even when the act is not reckless. For example, many courts have concluded that amputating the wrong limb is so obviously erroneous that an expert witness is not needed.¹⁴⁹ These acts may not result from reckless conduct, but they are clearly negligent. Many instances of negligence per se also constitute clear breaches of the standard of care.¹⁵⁰ There is no factual question that the plaintiff is a victim of wrongful conduct and has been physically and financially harmed as a result. In these cases, fee shifting is more consistent with the principle of compensation. Moreover, to the extent that the wrong is clear, such as with reckless conduct, greater compensation can be both a deterrent and a punitive measure.

This short discussion of the interplay between procedural optionality and tort law shows that procedural laws can be a vital factor in supporting and achieving the underlying policies of substantive entitlements. This idea is not new—procedural mechanisms can address policy objectives of substantive law. For example, shareholder derivative suits are subject to a particular pleading requirement that is designed to weed out meritless claims at the pleading stage.¹⁵¹ The novel idea proposed here is that the optionality of procedural rules can result in greater convergence of private and social interests. Accordingly, procedural flexibility and private ordering of public adjudication can potentially expand the means of achieving policy

¹⁴⁷ *Id.* at 160.

¹⁴⁸ *Id.* at 167.

¹⁴⁹ *Britt v. Taylor*, 852 So. 2d 1128, 1133 (La. Ct. App. 2003) (“Expert testimony is not required to meet the burden of proof in a medical malpractice case involving an obvious careless act from which a lay person can infer negligence, such as . . . amputation of the wrong limb”); *Young v. Key Pharms., Inc.*, 770 P.2d 182, 189 (Wash. 1989) (“Nonexpert testimony is sometimes admissible . . . [w]here the determination of negligence does not require technical medical expertise, such as the negligence of amputating the wrong limb”).

¹⁵⁰ See *supra* note 139.

¹⁵¹ FED. R. CIV. P. 23.1(b)(3) (requiring particularized pleading in shareholder derivative actions).

objectives in such fields as torts, securities, corporations, and civil rights, among others.

IV THE VALUE OF PRIVATE ORDERING IN PUBLIC ADJUDICATION

The preceding Parts of this Article proposed and explained a specific policy change to the litigation system and showed why that change would increase the number of meritorious suits brought, while simultaneously reducing the number of frivolous suits. While this policy suggestion is interesting in itself, it represents just one example of the value-creating potential of private ordering in public adjudication. This Part presents a broader theory of private ordering in public adjudication.

A. Private Resolution and Public Adjudication

We tend to think of legal disputes as typically resolved in one of two discrete, even if simultaneous, antipodal processes: private resolution or public adjudication. In economic terms, these processes constitute different pricing mechanisms.¹⁵² In private resolution, settlement occurs only if there is a clearing price. The price of the legal right is determined bilaterally, and thus settlement obviously requires the other party's cooperation and assent. If there is not a clearing price, the price-setting forum is public adjudication, and each party has the unilateral option to choose trial. There, the price of the legal right is judicially determined. The litigation system acts as a market in which information concerning dispute rights—which can be seen as economic assets and liabilities—is developed, traded, and assessed. Market participants resolve ambiguous claims through an informed “bet” on the litigated value of the disputed right.¹⁵³ The structure of the litigation system, however, limits such bets to a decision on settlement or trial, that is, a choice of a bilateral or judicial price-setting forum. Unfortunately, there is not a more sophisticated means of allowing parties to make informed, substantial bets on the

¹⁵² See generally Rhee, *supra* note 38 (analyzing “the complexities of the selection between settlement and litigation within the framework of economic efficiency”).

¹⁵³ See Rhee, *supra* note 33, at 226 (“In the absence of market pricing, each lawsuit is a market unto itself, and each party is forced to be a ‘market-maker’ for the other. . . . [T]he pricing of transactions within this micro-market determines settlement and litigation values.”); see also Jonathan T. Molot, *A Market in Litigation Risk*, 76 U. CHI. L. REV. (forthcoming 2009) (manuscript on file with the *New York University Law Review*) (arguing that litigation markets could be erected to price suits more efficiently).

quality of their case, independent of the other party's choice of settlement or trial.

Private resolution and public adjudication often occupy the same time and space, as is the case when settlement dialogue occurs during litigation, but they are parallel universes.¹⁵⁴ Although the parties may pursue parallel tracks, in the end they must select one. Either the parties or a court must ultimately determine the value of the legal rights in a dispute.

The distinct pricing mechanisms of private ordering and public adjudication also feature very different approaches to the process by which the price is determined. In private resolution, parties have free reign to rearrange their rights and obligations to their preference. Disputants can disregard violations of the law¹⁵⁵ or forgive them,¹⁵⁶ can weigh inadmissible evidence when considering the merits, or can compensate for attorney fees when the law does not so provide, just to give a few examples. Indeed, given the prominence of private resolution as the chosen pricing mechanism for dispute resolution, one can fairly suggest that the law "exists in the shadow of bargaining" in the sense that the policy objectives of public substantive laws are largely achieved in a private dispute resolution system.¹⁵⁷

¹⁵⁴ In the academic and policy debates, the two processes are often portrayed in stark black and white—in terms of good and bad—reflecting the polar nature of these processes. Scholars, judges, and policymakers overwhelmingly prefer private resolution. *See, e.g., In re Warner Commc'ns Sec. Litig.*, 618 F. Supp. 735, 740 (S.D.N.Y. 1985) (observing "the familiar axiom that a bad settlement is almost always better than a good trial"), *aff'd*, 798 F.2d 35 (2d Cir. 1986); Gross & Syverud, *supra* note 39, at 3 ("Trial is a disease, not generally fatal, but serious enough to be avoided at any reasonable cost." (footnote omitted)); Russell Korobkin & Chris Guthrie, *Psychological Barriers to Litigation Settlement: An Experimental Approach*, 93 MICH. L. REV. 107, 107–08 (1994) (noting that most commentators believe "trials represent mistakes—breakdowns in the bargaining process—that leave the litigants and society worse off than they would have been had settlement been reached"). Indeed, courts actively discourage disputants from opting for public adjudication. *See, e.g., Marek v. Chesny*, 473 U.S. 1, 10 (1985) ("In short, settlements rather than litigation will serve the interests of plaintiffs as well as defendants."); *G. Heileman Brewing Co. v. Joseph Oat Corp.*, 871 F.2d 648, 664 (7th Cir. 1989) ("Settling litigation is valuable, and courts should promote it."). Learned Hand's famous attitude toward litigation reflects the sentiment of many judges: "[A]s a litigant I should dread a lawsuit beyond almost anything short of sickness and death." Learned Hand, *The Deficiencies of Trials To Reach the Heart of the Matter*, Address Before the Association of the Bar of the City of New York (Nov. 17, 1921), in JAMES N. ROSENBERG ET AL., *LECTURES ON LEGAL TOPICS, 1921–1922*, at 89, 105 (1926).

¹⁵⁵ *See supra* note 41 (noting that many meritorious claims go unprosecuted).

¹⁵⁶ The law recognizes forgiveness as a form of private ordering. *See, e.g., CAL. EVID. CODE* § 1160 (West Supp. 2008) (providing evidentiary shields for benevolence and expressions of sympathy); MASS. GEN. LAWS ANN. ch. 233, § 23D (West 2000) (same); TEX. CIV. PRAC. & REM. CODE ANN. § 18.061 (Vernon 2008) (same).

¹⁵⁷ Rhee, *supra* note 41, at 127; *see also id.* at 168–69 (arguing that settlement reduces standard of care, as predicted by basic law and economics model of tort law).

In public adjudication, on the other hand, parties are granted far fewer opportunities to rearrange their rights and obligations. The court ensures strict adherence to the law, and the law is not subject to optional application or modification. Many rules are inflexibly mandated. If rules are subject to discretionary application, the court has the exclusive power to exercise it. For example, even when fee shifting is allowed, fee indemnity is not always an entitlement of the prevailing party; it is subject to judicial discretion.¹⁵⁸

There is almost no private choice of procedural laws in public adjudication and, conversely, no public mandating of procedural laws in private dispute resolution.¹⁵⁹ But why must private choice and public adjudication be viewed as essentially mutually exclusive?

Neither litigation nor settlement is a global panacea for the many problems of resolving private disputes. There has been voluminous commentary on the negative effects of litigation in our society, and so no further comment is needed here.¹⁶⁰ The generic, if not reflexive, indictment against litigation is unhelpful, as it too often rings of polemic. Certainly there are pockets of inefficiency and improvements to be made. For example, most would agree that low value cases are underlitigated and that frivolous cases are overlitigated. Absent the consideration of cost, litigation in both contexts would advance societal interests, but the existence of cost produces a divergence of private and social interests.

As for settlement, a commonly asserted shibboleth is that it is superior to public adjudication because it minimizes transaction costs.¹⁶¹ This is true, at least when cost is seen narrowly as cash outlays, instead of more broadly as incorporating risk-based valational adjustments.¹⁶² But positive settlement can also aggravate the problem of frivolous suits since ideally their settlement value would be zero; if settlement were not an option and all cases were required to proceed to trial, a frivolous suit would not be profitable. Moreover, private resolution of the dispute induced by the American rule may

¹⁵⁸ See, e.g., 29 U.S.C. § 1370(e)(1) (2006) (giving courts discretion to award fees to prevailing party); ARIZ. REV. STAT. ANN. § 12-341.01(A) (2003) (same).

¹⁵⁹ See *supra* Part I.A (discussing widespread imposition of inalienable procedural rules in public adjudication).

¹⁶⁰ Cf. Rhee, *supra* note 38, at 621 n.2 (citing sources regarding preferability of settlement to litigation in court).

¹⁶¹ See Larry Kramer, *Consent Decrees and the Rights of Third Parties*, 87 MICH. L. REV. 321, 327 (1988) ("Settlement is more efficient for the parties, giving them more of what they hoped to gain at less cost."); David M. Trubek et al., *The Costs of Ordinary Litigation*, 31 UCLA L. REV. 72, 122 (1983) ("[B]argaining and settlement are the prevalent and, for plaintiffs, perhaps the most cost-effective activity that occurs when cases are filed.").

¹⁶² See *supra* note 33.

reduce the value of many low value cases effectively to zero, if transaction costs exceed the expected value. The degree of positive value depends on the defendant's aversion to risk and the plaintiff's credible threat to pursue public adjudication given the transaction costs (ironically, the same calculation as that of the plaintiff in a frivolous suit).

Clearly, without transaction costs, public adjudication would remedy the problem of frivolous suits and unprosecuted low value cases. While settlement is always bilateral, thus giving rise to the inference of mutual utility maximization, it does not maximize enforcement and compliance at the lowest cost in many cases. Settlement, so often seen as an optimal solution to the transaction cost problem of litigation, is a blunt instrument, and it ignores important problems through a single-minded global concern for monetary cost savings.¹⁶³

B. The Need for Private Ordering in Public Adjudication

This Article rejects the view that private ordering and public adjudication are mutually exclusive, for there is no compelling theoretical basis to support this wall. By rejecting the old axiom, we can construct alternative processes for dispute resolution with greater flexibility. Once public adjudication is selected, many decisions affecting the allocation of value, cost, and risk should be placed in the hands of the parties, who have the best information about a claim. Greater efficiencies in the realm of public adjudication result when the function of the court is reduced to the price-setting body of last resort. It does not follow from the fact that a court—rather than the parties—sets the price of the legal right in public adjudication that the court must also inflexibly set the *processes* by which that price is determined.

Outside of the price-setting function itself, the law should promote private ordering even within public adjudication. It is true that in this regard, courts routinely (perhaps aggressively) push settlement¹⁶⁴ and that legislatures have enacted rules to promote it.¹⁶⁵ But

¹⁶³ Accordingly, some scholars, swimming upstream perhaps, have criticized the view that settlement is normatively superior to litigation. See Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1075 (1984) ("I do not believe that settlement as a generic practice is preferable to judgment or should be institutionalized on a wholesale and indiscriminate basis."); Rhee, *supra* note 38, at 625 ("The axiomatic belief that settlement is inherently superior is regrettable, for it is far from clear that as a general rule, settlement—more precisely, settlement independent of litigation—provides the least costly economic transaction in contested actions.").

¹⁶⁴ See cases cited *supra* note 154.

¹⁶⁵ See, e.g., FED. R. CIV. P. 68 (establishing offer of judgment procedure and presenting potential negative consequences of rejecting settlement offers).

it is a mistake to equate the private ordering of disputes with settlement. The former is a process; the latter is an outcome. Private ordering of a dispute is achieved when the parties are allowed to rearrange the procedural rules governing public adjudication of the dispute so as to reach an efficient result at the lowest individual and social cost. Because procedural rules can impose significant costs and risks on the parties, some rules should be subject to reordering by the parties.

The idea of optional procedural rules is informed by Coase's work. In *The Problem of Social Cost*, Coase showed that if transaction costs were zero, parties would rearrange their rights in a way that maximizes efficiency irrespective of the initial assignment of the rights.¹⁶⁶ That is, in the context of economic production, the law is irrelevant prior to the existence of transaction costs.¹⁶⁷ Once transaction costs are considered, however, the rearrangement of rights only occurs if the increase in production exceeds the cost of bringing about this reordering.¹⁶⁸ Coase explained:

In these conditions the initial delimitation of legal rights does have an effect on the efficiency with which the economic system operates. One arrangement of rights may bring about a greater value of production than any other. But unless this is the arrangement of rights established by the legal system, the costs of reaching the same result by altering and combining rights through the market may be so great that this optimal arrangement of rights, and the greater value of production which it would bring, may never be achieved.¹⁶⁹

The broad lesson here is that the law should initially assign rights in a way that reflects the hypothetical bargain of the parties but *should leave the parties the option of rearranging those rights through private ordering*.¹⁷⁰ Coase argued that government interference with private ordering can be costly:

¹⁶⁶ Coase, *supra* note 8, at 2-7, 15.

¹⁶⁷ In his parable of the farmer and cattle raiser, Coase showed that, in the absence of transaction costs of bargaining and in conditions of perfect competition, parties who are engaged in mutually harmful activity would achieve a bargain that maximizes economic production irrespective of the initial assignment of the entitlement. *Id.* at 2-6. Of course, there is a wealth effect of the initial assignment of rights to the parties involved in the activity.

¹⁶⁸ *Id.* at 15-16.

¹⁶⁹ *Id.* at 16.

¹⁷⁰ For years, courts have used the analytic heuristic of an *ex ante* bargain or deliberation to determine the most efficient rule of law. See, e.g., *Rodi Yachts, Inc. v. Nat'l Marine, Inc.*, 984 F.2d 880, 888-89 (7th Cir. 1993) (Posner, J.) ("[T]he market itself fixes a standard of care that reflects the preferences of potential victims as well as of potential injurers and then the principal function of tort law, it could be argued, is to protect customers' reasonable expectations [of] . . . the standard of care customary in the industry" (citing *U.S. Fid. & Guar. Co. v. Jadranska Slobodna Plovidba*, 683 F.2d 1022, 1029 (7th Cir. 1982)));

It is clear that the government has powers which might enable it to get some things done at a lower cost than could a private organisation (or at any rate one without special governmental powers). But the governmental administrative machine is not itself costless. It can, in fact, on occasion be extremely costly. . . . But equally there is no reason why, on occasion, such governmental administrative regulation should not lead to an improvement in economic efficiency.¹⁷¹

While Coase's analysis above regards the context of the government's substantive rulemaking powers, specifically the law of nuisance and property, its force applies equally to procedural entitlements and obligations, because they, like substantive laws, greatly influence the efficiency of the dispute outcome.

These insights show why neither the "settlement-versus-litigation" nor the "best rule" debate produces a complete solution to the many different but interconnected problems of resolving private disputes. In each case, the bargain of the parties may be different, depending on the unique set of facts, laws, circumstances, and risks.¹⁷² Broad solutions inherently conflict with the nature of disparate problems.

Procedural rules substantially determine the level of risk and the amount of transaction costs each party must assume. Many rules are flexible and subject to judicial discretion, and many are not.¹⁷³ The authority to invoke, reject, or reformulate many of these rules never resides with the parties, for it has always been assumed that the sovereign generally mandates the law, often in an inflexible manner. The issue is obvious. Given the unique circumstances of each case, can a variable legal standard, resulting from each party's unilateral choice rather than a fiat of the court or legislature, minimize transaction costs while maintaining important jurisprudential considerations of fairness and predictability?

No single rule of procedure can be a global panacea. As shown in the above discussion, both the American and English rules have pros and cons, the weight of which depend on the individual circum-

Bamford v. Turnley, (1862) 122 Eng. Rep. 27, 33 (Exch.), *rev'd* (1860) 122 Eng. Rep. 25 (Q.B.) (analyzing nuisance case from hypothetical decision of individual owner of properties in question); *see also* Richard A. Epstein, *A Theory of Strict Liability*, 2 J. LEGAL STUD. 151, 157–58 (1973) (arguing that *Vincent v. Lake Erie Transportation Co.*, 124 N.W. 221 (Minn. 1910), applied *ex ante* bargaining analysis to determine whether there should be liability for damages in case of private necessity).

¹⁷¹ Coase, *supra* note 8, at 17–18.

¹⁷² Nor is arbitration the answer. It requires an *ex ante* contract or a legislative mandate, and it still is founded on adjudication by a third party.

¹⁷³ *See supra* Part I.A (discussing widespread imposition of inalienable procedural rules in public adjudication).

stances.¹⁷⁴ Many rules are currently legal constants and not subject to modification at all, such as the standard of proof. The search for the best rule or approach toward broad problems is inconclusive because broad rules have an inherently indiscriminate quality that is sometimes inimical to contextual problem solving.

Some scholars have taken an alternative tack and considered the permutations of different procedural aspects of dispute resolution. The effort of Lucian Bebchuk and Howard Chang is notable.¹⁷⁵ They sought to combine attorney fee rules with the margin of victory, and a formal model shows how this proposal would work under various conditions of information uncertainty, information asymmetry, and the proposal's implications on Rule 11. The court determines the margin of victory, and the size of this margin may be subject to policy considerations, such as the deterrence of frivolous suits, incentivization of public interest litigation, and so on. Thus, Bebchuk and Chang connect attorney fees to the quality of a party's win or loss, which is the essential Rule 11 calculus.¹⁷⁶

While the theory of connecting fee indemnity to case merit is sound, their proposal is problematic. From an instrumental view, defining the precise point of sufficient margin is difficult, and absent a clear definitional standard, we would expect significant variance among rulings from courts. At issue is unpredictability, which always raises costs on the parties and the litigation system. The courts not only must undertake a laborious common law process of rulemaking on individual cases but they must also determine whether classes of lawsuits and causes of action should be subject to a margin of victory scaling.¹⁷⁷ The process of case-by-case, class-by-class review may result in instability of the law, significant variations among fora and jurisdictions, and greater uncertainty of outcome.

Judicial control of the fee rule is problematic for two reasons. First, since the rule presumably applies uniformly to the entire spectrum of litigated cases, parties may be incentivized to litigate more fiercely on the hope that they can prevail by a wide margin. Most cases probably cannot meet the standard regardless of the parties' increased efforts, but because there may be great uncertainty, the par-

¹⁷⁴ See *supra* notes 15–23 (discussing various fee rules).

¹⁷⁵ Bebchuk & Chang, *supra* note 57, at 371–72.

¹⁷⁶ *Id.* at 374.

¹⁷⁷ There are instances in which courts disfavor or view with suspicion certain kinds of cases for various policy reasons. In these matters, the favored procedural device is the particularized pleading. See, e.g., FED. R. CIV. P. 23.1(b)(3) (providing for particularized pleading in shareholder derivative suits); DEL. CH. R. 23.1 (same). See generally Christopher M. Fairman, *Heightened Pleading*, 81 TEX. L. REV. 551, 574–77 (2002) (observing that heightened pleading developed in part to deter frivolous claims).

ties' efforts may spiral upward, since they may be incentivized either to seek a wider margin of victory or to avoid one. This effect undermines the goal of minimizing the cost of litigation. Private choice of the rules regarding fee shifting and standards of proof may promote self-regulation of litigation efforts better than broadly applicable rules, particularly where there is significant uncertainty.

Second, during litigation and up to judgment, the parties are uncertain as to which fee rule will govern, since the selection of the rule is contingent on the "margin of victory" at the litigation's outcome. This uncertainty may be sufficient to discourage a plaintiff with a low value case and a defendant facing a frivolous action. Without greater clarity on the fee rules earlier in the litigation, these parties might not gamble significant and certain cost expenditures for the slight possibility of prevailing with fee recovery. Instead, the low cost options of not prosecuting the dispute or choosing low value settlement, respectively, may be better and safer investment choices.¹⁷⁸

The theory of procedural optionality allows solutions to the choice-of-fee-rule problem like the proposal presented in the preceding Parts. As the previous sections explained, we have compartmentalized our thinking into discrete policy debates on the best dispute-resolution process and the best rule of law.¹⁷⁹ The Bebachuk and Chang analysis is a laudable effort to break this habit of thought, but it falls into the same pattern of working from the axiomatic premise that private ordering can be useful only in the context of private dispute resolution. The drawbacks of their approach are the inevitable consequence of the axiom that public adjudication requires an inflexible, judge-centric imposition of legal rules that allocates value, cost, and risk upon the parties who cannot then rearrange these important matters within the context of public adjudication.

¹⁷⁸ See Herbert M. Kritzer, *The English Rule*, A.B.A. J., Nov. 1992, at 57 ("If the amount at stake were \$10,000 or \$25,000, most middle-income individuals still would be reluctant to put \$5,000 to \$10,000 on the line to pursue even a strong case.").

¹⁷⁹ See John W. Wade, *On Frivolous Litigation: A Study of Tort Liability and Procedural Sanctions*, 14 HOFSTRA L. REV. 433, 494 (1986) ("What is needed is a conscious effort to coordinate the procedural and tort law, so that the trial can be handled as simply as possible without the interference of frivolous litigation abuses and without inconsistent results in the original and second trials."). Additionally, legal analysis tends to compartmentalize rules of law according to their traditional taxonomy. See Robert J. Rhee, *A Principled Solution for Negligent Infliction of Emotional Distress Claims*, 36 ARIZ. ST. L.J. 805, 865 (2004) ("Judicial and scholarly analyses have tended to compartmentalize rules of liability and damage, and have created an intellectual inertia, hindering critical analysis.").

CONCLUSION

The idea that litigants can unilaterally elect the procedural rules to be applied in their cases is intellectually jarring only if one sees a substantial conceptual barrier between private resolution and public adjudication. The legitimacy of this barrier is universally accepted, which is unfortunate since a greater combination of private ordering and public adjudication could have highly beneficial effects. This acceptance also creates contradictions between positive observations and normative ideals. It is obvious that if parties could resolve their disputes privately, they would not resort to costly adjudicatory processes. It is easy to conclude that public adjudication is a result of error or bad judgment. But it is just as easy to conclude that in a world of ambiguous facts and laws, reciprocity of harm, and subjective projections of inherently imperfect human decisionmaking, some disputes cannot be resolved without neutral third-party intervention. Once adjudication is selected, the parties have committed to undertaking greater risks and costs. The risks and costs of litigation are well known, and they are fixed in the sense that the rules of law imposing them generally cannot be varied through the parties' unilateral election.

This Article advances the theory that attributes of both private ordering and public adjudication can be blended to provide more efficient dispute resolutions. This proposition flows from Coase's analysis of social cost. The problem with the application of inalienable rules of law is that each rule works well in some circumstances but not in others. Sometimes these rules are inflexible due to the legitimate policy considerations of fairness and predictability, and sometimes discretionary application requires substantial litigation. But fixed rules of law can impart costs on the litigation system as well. Even in the context of public adjudication, private parties are in the best position to rearrange ambiguous rights so long as strategic behavior is minimized through the bonding of good faith.

The benefits of procedural optionality can be seen in this Article's election scheme regarding the attorney fee rules and standards of proof. This proposal specifically addresses the social problems of frivolous litigation and low value cases. The public interest is advanced when frivolous suits are dismissed and low value cases are prosecuted. Yet the cost structure of litigation promotes behavior in direct opposition to this social interest. This can result in suboptimal enforcement of meritorious claims and in costly expenditures to avoid defending frivolous claims. This Article shows that private ordering of the rules of law can remedy this problem. The parties

themselves, through their unilateral choice of fee shifting and concomitant bonding of their good faith belief in the clear merit of their claim by assuming a higher standard of proof, can provide appropriate incentives and disincentives in a way that yields greater social efficiency than under our current regime consisting mostly of inflexibly applied rules.